

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2004, 10:57:31 ; Search time 6273 Seconds

(without alignments)
10799.484 Million cell updates/sec

Title: US-09-992-901-1

Perfect score: 1563
Sequence: 1 atggaggaagaagtagcag.....tgaccatcatgaggtatga 1563

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GeneBtbl:
1: gb_da:
2: gb_hg:
3: gb_in:
4: gb_cm:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vl:
15: em_da:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_cm:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vl:
30: em_hg_hum:
31: em_hg_inv:
32: em_hg_other:
33: em_hg_mus:
34: em_hg_pin:
35: em_hg_rtd:
36: em_hg_mam:
37: em_hg_vrt:
38: em_sy:
39: em_hgo_hum:
40: em_hgo_mus:
41: em_hgo_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1559.8	99.8	1563	BT010564	BT010564 Arabidops
2	873.6	55.9	9167	AC003105	AC003105 Arabidops
3	400.6	25.6	2151	AK106672	AK106672 Oryza sat
4	318.8	20.4	1539	AX467615	AX467615 Sequence
5	318.8	20.4	1539	BT003388	BT003388 Arabidops
6	318.8	20.4	1786	AF407303	AF407303 Arabidops
7	291.6	18.7	1548	AF446870	AF446870 Arabidops
8	290.2	18.6	1784	AY052208	AY052208 Arabidops
9	290.2	18.6	1860	AK120018	AK120018 Oryza sat
10	289.6	18.5	1854	CTRCPP450P	CTRCPP450P
11	281.2	18.0	1633	CTRP450B	CTRP450B
12	276.4	17.7	1570	AY067748	AY067748 Arabidops
13	276.4	17.7	1692	AY050827	AY050827 Arabidops
14	273.2	17.5	1572	AB047400	AB047400 Oryza sat
15	273.2	17.5	1903	AK101545	AK101545 Oryza sat
16	272.2	17.4	1813	AB067685	AB067685 Solanum t
17	271.6	17.4	2441	AK105255	AK105255 Oryza sat
18	269.2	17.2	1539	AX467616	AX467616 Sequence
19	269.2	17.2	1539	BT001008	BT001008 Arabidops
20	269.2	17.2	1697	AF367347	AF367347 Arabidops
21	265.8	17.0	1570	BT002745	BT002745 Arabidops
22	265.8	17.0	1618	AY065060	AY065060 Arabidops
23	265.4	17.0	1833	AX467614	AX467614 Sequence
24	259.2	16.6	2227	AF321870	AF321870 Lolium ri
25	257.6	16.5	1666	AF321866	AF321866 Lolium ri
26	257.6	16.5	1739	AF321869	AF321869 Lolium ri
27	253.2	16.2	1821	AF321868	AF321868 Lolium ri
28	253.2	16.2	1865	AF321867	AF321867 Lolium ri
29	250.4	16.0	1620	AY437442	AY437442 Triticum
30	248	15.9	1767	AF321870	AF321870 Lolium ri
31	246.6	15.8	1963	AK101750	AK101750 Oryza sat
32	246.2	15.8	1811	AK101667	AK101667 Oryza sat
33	245.6	15.7	2098	AK073238	AK073238 Oryza sat
34	244.4	15.6	1587	AB038597	AB038597 Oryza sat
35	241.2	15.4	1557	AX050397	AX050397 Sequence
36	236.4	15.1	1734	AF249329	AF249329 Lycopersi
37	235.8	15.1	1849	AK101513	AK101513 Oryza sat
38	235.8	15.1	2167	AK120700	AK120700 Oryza sat
39	230.6	14.8	1560	AF091688	AF091688 Arabidops
40	230.6	14.8	1855	AF367329	AF367329 Arabidops
41	220.8	14.1	67349	AP003752	AP003752 Oryza sat
42	220.8	14.1	101607	AP004334	AP004334 Oryza sat
43	216.4	13.8	1901	AK074025	AK074025 Oryza sat
44	216.2	13.8	1838	AK072220	AK072220 Oryza sat
45	207.6	13.3	1815	AK107349	AK107349 Oryza sat

ALIGNMENTS

RESULT 1
BT010564 1563 bp mRNA linear PLN 30-AUG-2003
LOCUS Arabidopsis thaliana At2g26710 gene, complete cds.
DEFINITION Arabidopsis thaliana At2g26710 gene, complete cds.
ACCESSION BT010564.1 GI:34365750
VERSION BT010564.1
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1563)
CHECK, R., Chen, H., Kim, C.J., Shin, P., Bowser, L., Carninci, P.,
Dale, J.M., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A.,

TITLE
Journal
REFERENCE
AUTHORS

Karlín-Neumann, G., Kawai, J., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.
Arabidopsis ORF clones
Unpublished
2 (bases 1 to 1563)

TITLE
Journal
REFERENCE
AUTHORS

Chen, R., Chen, H., Kim, C.J., Shim, P., Bowser, L., Carninci, P., Dale, J.M., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlín-Neumann, G., Kawai, J., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.
Direct Submission
Submitted (30-AUG-2003) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT

The discrepancy does not affect the protein sequence.
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y., and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Chen, R., Chen, H., Kim, C.J., Shim, P., Bowser, L., Chan, M.M., Chang, C.K., Dale, J.M., Hsuan, V.M., Jones, T., Karlín-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Chen, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES
SOURCE

Location/Qualifiers
1..1563
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="U61967"
/ecotype="Columbia"
/note="This clone is in PUNI 51"
1..1563
/note="cytochrome P450"
/codon_start=1
/product="At2g26710"
/protein_id="AA065187.1"
/db_xref="GI:34365751"
/translation="MEBESSWFIPVILVSLVILVKSGLIMBPRKIEHFSK QGIRPPIFFIGVKEIVGMKASHPMPRHNIIPVLSFYHMKRIVATLW FGPTFLVADPDLREIFSKSEFEKSEHNPVYKQLESDGLSLGKRMHRRILIS PTFTMENIKLIVPVILKSVTDVWDKMSDKSENGEVVDVIEWFOLETDVSRIRAG SSYEDGAVFRLAQOMLCAEAFQKVIPEGAFPTFRTNLSKWDKEIRKSLKLI EERRQNAIDGEGCEKPAADLGLMIQAKNTVODIYEECKSPFPAKQTSMLLT WTTLLSMHPENOKAKDEVLRVGSVDVPTKDHVVKLTLSMINSLSRLTPIVAT TRKRSYVKAGYKIPGCTELLIPITAVHNDQIWGNDVNEPAPFADGVRAKHP VGRIPFGVATCIGONLAILQAKLTIAWIKQFTFLAFTYQHPATVLMILYPOHGA PTFPRRLTNHED"
1356
/note="compared to genome"
/replac="c"

ORIGIN

Query Match 99.8%; Score 1559.8; DB 8; Length 1563;
Best Local Similarity 99.9%; Freq. No. 0;
Matches 1561; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGAAAGAAAGTAGCAGCTGATTCATTCAGAAAGTCTGTCGTGTAATCTTA 60
DB 1 ATGAGAGAAAGAAAGTAGCAGCTGATTCATTCAGAAAGTCTGTCGTGTAATCTTA 60
QY 61 AGCTCTGTAATAGTAAGAGGTAATGTCCTGTTATGTTGAGACCAAGAAAGATTGAAGA 120
DB 61 AGCTCTGTAATAGTAAGAGGTAATGTCCTGTTATGTTGAGACCAAGAAAGATTGAAGA 120
QY 121 CATTCCTTAACAAGGAATTCAGAGTCCTCCTTATCATTTCTTCATCCGAAATGTAA 180
DB 121 CATTCCTTAACAAGGAATTCAGAGTCCTCCTTATCATTTCTTCATCCGAAATGTAA 180
QY 181 GAACCTGTGGAATGATGCTTAAAGCTTCTTCATCTATGCTTCTCTCACATATT 240
DB 181 GAACCTGTGGAATGATGCTTAAAGCTTCTTCATCTATGCTTCTCTCACATATT 240
QY 241 CTTCCTAGAGTCTCTCTTTTACCATCAGGAGAAATCTACGGTGTACATTTCTG 300
DB 241 CTTCCTAGAGTCTCTCTTTTACCATCAGGAGAAATCTACGGTGTACATTTCTG 300
QY 301 GTTTGTTCCGTCCTCACTTCCGTTTACCGGTACCGATCTGTATTTGATCAGAGATC 360
DB 301 GTTTGTTCCGTCCTCACTTCCGTTTACCGGTACCGATCTGTATTTGATCAGAGATC 360
QY 361 TTCTCTAAGTCTGATCTTCTACGAGAAATGAACTCACCCCTTTGTTAAACAATTGAA 420
DB 361 TTCTCTAAGTCTGATCTTCTACGAGAAATGAACTCACCCCTTTGTTAAACAATTGAA 420
QY 421 GGCAGATGACCTACTAGTCTCAAGGTGAAAGAGGGCTCATCTCGAAATATCTAGC 480
DB 421 GGCAGATGACCTACTAGTCTCAAGGTGAAAGAGGGCTCATCTCGAAATATCTAGC 480
QY 481 CCTACTTTCAATATGAGAAATCTTAAAGTCTTGTACCAAGTTGTGGAAGAGTGAAT 540
DB 481 CCTACTTTCAATATGAGAAATCTTAAAGTCTTGTACCAAGTTGTGGAAGAGTGAAT 540
QY 541 GATATGATGATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 GATATGATGATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 TATGATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 TATGATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 GCTTTTCAAAAAGTCTTCAATTCCTGCTATGATTTTTCGCAACAAGGGAATTGAG 780
DB 721 GCTTTTCAAAAAGTCTTCAATTCCTGCTATGATTTTTCGCAACAAGGGAATTGAG 780
QY 781 TCTCGAAGTGTAGCAAGAGATTAAGAGTGTGTTGAAGCTATGAGCGCGGAGAGA 840
DB 781 TCTCGAAGTGTAGCAAGAGATTAAGAGTGTGTTGAAGCTATGAGCGCGGAGAGA 840
QY 841 CAAGAAGCTATATGATGAGAGAGGGAAGATTAAGAGCGCGCGGAGAGATTTGTTG 900
DB 841 CAAGAAGCTATATGATGAGAGAGGGAAGATTAAGAGCGCGCGGAGAGATTTGTTG 900
QY 901 GATTAATGATTAAGGCAAGAAATGAGAGGTTCAAGGACATGTGAGAGAGTGAAGC 960
DB 901 GATTAATGATTAAGGCAAGAAATGAGAGGTTCAAGGACATGTGAGAGAGTGAAGC 960
QY 961 TTTTCTTCCGCGGAGAAAGAGCACTTCTATCTGCTGAGTGAAGCAATCTTGCTA 1020
DB 961 TTTTCTTCCGCGGAGAAAGAGCACTTCTATCTGCTGAGTGAAGCAATCTTGCTA 1020
QY 1021 TCCATGACCCCGAGAGTGGCAGGCAAGACCGATGAGAGTCTCGAGGCTCGAGCTCA 1080
DB 1021 TCCATGACCCCGAGAGTGGCAGGCAAGACCGATGAGAGTCTCGAGGCTCGAGCTCA 1080
QY 1081 CGTATGTCCTTACCAAGACATGTCGTTAAGCTTAAACGTTAGATGATCTTGAAAC 1140

Db		52416	GCTGATAGACGGCCGGAGACAACAAACCCTTAATGTGAAGAAGCGAAGANTGTAAAGAGCC	52475
Qy		882	GCGCGCAAGAGATTGTGGAGATTAAATGATTCAAGCAAAGATGTGACGGTTCAAGACAT	941
Db		52476	GCGCGCAAGAGATTGTGGAGATTAAATGATTCAAGCGCAAGAAATGTGACGGTTCAGACAT	52535
Qy		942	TGTGGAGAGAGTAAAAAGCTTTTCTTCGCGCGGAAAAGACAACCTTCAATCTGTAC	1001
Db		52536	TGTGGAGAGAGTAAAAAGCTTTTCTTCGCGCGGAAAAGACAACCTTCAATCTGTAC	52595
Qy		1002	GTGACGACCATCTTGCTATCCATGCACCCGGAAGTGGCAGGCCCAAGCACGTGATGAGT	1061
Db		52596	GTGACGACCATCTTGCTATCCATGCACCCGGAAGTGGCAGGCCCAAGCACGTGATGAGT	52655
Qy		1062	CCTCAGGCTGTGGGGCTCAGGTATATCCCTCCACAGACCAATGTGCTTAAGCTAAAG	1121
Db		52656	CCTCAGGCTGTGGGGCTCAGGTATATCCCTCCACAGACCAATGTGCTTAAGCTAAAG	52715
Qy		1122	-----	1121
Db		52716	GGTGGCTTCTCTTCTGTCTTGGCTTGTCTTTCATATTTGAAATATATCATTA	52775
Qy		1122	-----	1121
Db		52776	ATTACAGATAAAGGCGAATTTGTGATAGACAGTTCCGTTACGTTAATTATGTCTTT	52835
Qy		1122	-----	1121
Db		52836	TTGTGCGCTAATCTAATAATTATATATGTGCTAATAATCTACATCCATCAAGGCAAGT	52895
Qy		1122	-----	1121
Db		52895	TCAAGGCAAAAGGATATATGATCATAGAAAGATATGTGATGTTGTTAACCTCTTCA	52955
Qy		1122	-----GTTGATATGATCTTGAACGAGCTCTTAAAGTTGATTCACCAATAGAGT	1173
Db		52956	AATTTGCAATTTGATGATGATCTTGAACGAGCTCTTAAAGTTGATTCACCAATAGAGT	53015
Qy		1174	ACGATTCGACGCGCTAAATCGATGTGAAGCTAGAGGGTACMAAATCCCATGTGACAG	1233
Db		53016	ACGATTCGACGCGCTAAATCGATGTGAAGCTAGAGGGTACMAAATCCCATGTGACAG	53075
Qy		1234	GAGCTTCTAATCCCAATCATAGAGGCTCATCATGACCAAGCCATTTGGGGTATAGACGT	1293
Db		53076	GAGCTTCTAATCCCAATCATAGAGGCTCATCATGACCAAGCCATTTGGGGTATAGACGT	53135
Qy		1294	AACGAATTCATCCAGCTCGGTTTGGAGTAGAGTCCGCGTCTCCAAACACCCCGTT	1353
Db		53136	AACGAATTCATCCAGCTCGGTTTGGAGTAGAGTCCGCGTCTCCAAACACCCCGTT	53195
Qy		1354	GAGCTTCATACCGTTTGGCTTCGGATTCGATCATGCTATGTTGTGAAATCTTGCTATACT	1413
Db		53196	GAGCTTCATACCGTTTGGCTTCGGATTCGATCATGCTATGTTGTGAAATCTTGCTATACT	53255
Qy		1414	CAGGCCAAATTTGACACTCGCGCTGATATGATCCAAGCTTCACTTCACTTGGCTCTACT	1473
Db		53256	CAGGCCAAATTTGACACTCGCGCTGATATGATCCAAGCTTCACTTCACTTGGCTCTACT	53315
Qy		1474	TATCAGCATGACCTTACCGTCTTATGTGCTTATCTCAACATGGTGCACCAATCAC	1533
Db		53316	TATCAGCATGACCTTACCGTCTTATGTGCTTATCTCAACATGGTGCACCAATCAC	53375
Qy		1534	TTCCGAGATTTGACCAATCAATGAGAGATTGA	1563
Db		53376	TTCCGAGATTTGACCAATCAATGAGAGATTGA	53405

RESULT 3
AKI06672

Locus Oryza sativa (japonica cultivar-group) cDNA clone:002-114-A02, full insert sequence.

ACCESSION
AKI06672

VERSION
AKI06672.1 GI:32991881

KEYWORDS
FLI cDNA, oligo capping.

SOURCE
Oryza sativa (japonica cultivar-group)

ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
1
The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team:
Kikuchi, S., Satch, K., Nagata, T., Kawasashira, N., Doi, K., Kikuchi, S., Satch, K., Nagata, T., Kawasashira, N., Doi, K., Kishimoto, N., Yaaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Chentouk, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Otsomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kuroski, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Naikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikiura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., K. RIKEN, Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., Kawai, J., Carrincci, P., Adachi, J., Aizawa, K., Arikawa, T., Fukuda, S., Hara, A., Hashidume, M., Hayatsu, N., Imotani, K., Ishi, Y., Itoh, M., Kagawa, I., Kondo, S., Komoto, H., Miyazaki, A., Otsomo, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)

TITLE
JOURNAL
MEDLINE
PUBMED
12869764
2 (bases 1 to 2151)

REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akinura, T., Arikawa, T., Carrincci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, M., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imanura, K., Imotani, K., Ishibiki, J., Ishi, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kikuchi, S., Kishikawa, Hirozane, T., Kishimoto, N., Kobayashi, M., Kikuchi, S., Kishikawa, Hirozane, T., Kishimoto, N., Kobayashi, M., Kikuchi, S., Kishikawa, Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Komoto, H., Kouda, M., Koyama, S., Kurikawa, C., Kuroski, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Naikawa, R., Nikiura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Chentouk, K., Ooka, M., Ooka, H., Otsomo, N., Ota, Y., Otsomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakane, N., Sano, H., Sasaki, D., Sato, K., Satch, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami, Takada, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, Y., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.
Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica rice.
URL: <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satch, K., Nagata, T., Kawasashira, N., Doi, K., Kishimoto, N., Yaaki, J., Kishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Chentouk, K., Shishiki, T., and Yamamoto, M.
PAIS Genome Sequencing & Analysis Group: Otsomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kuroski, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Naikawa, R., Nikiura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hara, K.,
Hara, A., Hashizume, W., Hayashida, K., Hayashita, N., Hiramoto, K.,
Hirooka, T., Hori, F., Iida, J., Imamura, K., Imoto, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Komori, H., Kouda, M.,
Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N.,
Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami, T., Tanaka, A., Takahashi, F.,
Takaku, A., Tanaka, T., Tomaru, A., Toyota, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

FEATURES

source
1. 2151
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultiyar="Nidponbare"
/db_xref="taxon:39947"
/clone="002-114-A02"

ORIGIN

Query Match 25.6%; Score 400.6; DB 8; Length 2151;
Best Local Similarity 58.0%; Pred. No. 1e-103;
Matches 861; Conservative 0; Mismatches 569; Indels 69; Gaps 7;
90 GTTATGCTGAGACCAAGAAAGATTGAAGACATTCTTAACAAGAAATTCAGGCTC 145
175 GGTGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 234
150 TCTTATCATTTCTTCAATCGGAATGTTAAGAATCTTGATGATGATGATGATGATG 209
235 GCGCTACCGGTTCTCGTGGGCTGCGTGAAGAGATGAGGCTGATGAGGAGGAGC 294
210 TTCTGATCTTATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 263
295 GCGGAGCCGATGCG 354
264 CCATCACTGAGAGAAATCTACGCTGCTACATTTCTGTTGGTTCGCTCAACTTTCG 323
355 CCATCACTGAGAGAAATCTACGCTGCTACATTTCTGTTGGTTCGCTCAACTTTCG 414
334 GTTAAAGCTTACCGATCTGATTTGATCAAGAGATCTTCTTAAG---TCTGATTTCT 380
415 GCTACGCTGCGGAGCCCGGAGATGCTGCGGAGATCTTCTCAAGCGCGCGGAGCGCT 474
391 CGAGAAAGATGAGCTACCTCTTTGTTAAACAATGTAAGCGATGAGCTACTTACTCT 440
475 CGACCGCTACGAGGCGGACCCCGGCTGCGGAGCTGAGGAGGAGGAGGAGGAGGAG 534
441 CAAGGTGAAAAATGGGCTCATCATGAAAAATCATTAAGCTTCTTCAATGAGAA 500
535 CCAAGGCGACAAAGGCTCAACACGCGCGCTCTCAACCCCGGCTTCTTACCCGACAA 594
501 TCTTAAGTTCTTGTACAGTTGTTGTAAGAGTGTGATGATGATGATGATGATGATG 560
595 CCTCAACCGGCTGCTGCGGACGTCGACAGGTGCGTGGCGGCGGCGGCGGAGAGTGG 654
561 CGATTA---AGTTATCAAAAAAGGTTGAGTGTGAGTGTGATGATGATGATGATGAT 617
655 CGCATGCGCTGCG 714
618 TTGACGTGAAGATTTTATGAGAACAGCTTTTGAAGTATGATGATGATGATGATGATG 677
715 GGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 774
678 AGTTTTCGACTTCAAGCTCAACAAAGCTTCTTGTGTAAGCTTTTCAAAAGTCTT 737
775 CGTGTTCGCTTGAAGGCGCGCTCATGAGGCTTCTGAGGCTTCTGAGGAGGAGGCT 834
738 CATTCCTGCTATGATTTTCTCGACAAAGAGGAAATTGAAGTCTCGGAGATTGACAA 797

835 CCGTCCCGGAGTACAGGTTCTGTCCGACCAAGAAAGAACAGGATGTCGTGGGCGCTGAGAC 894
798 GGAAGTAAAGAGTCTGTTTGAAGCTGATGAGCGCGGAGACAAAGCGTATGATGAG 857
895 GGAAGTAAAGAGTCTGTTTGAAGCTGATGAGCGCGGAGACAAAGCGTATGATGAG 954
858 AAGAGAGG-----AAGATGTAAGAGCGCGCGGAGAGGATTTGTT 899
955 GGAAGTAAAGAGTCTGTTTGAAGCTGATGAGCGCGGAGACAAAGCGTATGATGAG 1014
900 GGAATTAATGATTA-----GCAAGAAATGAGCGCTTCAAGCAATTTGAGAG 950
1015 GGGGCTCATGATCAATGCGCGGCTGAGACAGACAGATGCGGTGAGAGACATGTTGAGAG 1074
951 GTGTAAAGCTTTTCTTCCGCGGAGAAACAGACAACTTCTMATCTGCTGACGTGACGAC 1010
1075 GTGCAAGACCTTCTTCTTCCGCGGAGAGACAGACCAACCACTGCTCACTGCGGAC 1134
1011 CATCTTCTATTCATGACCGCGGAGTGGAGCGGACCAAGACGATGAGGCTCTGAGGT 1070
1135 CGTGTCTGCTCGCATGACCGCGGCTGAGACAGACAGCGCGCGCGCGCGCGCGCGCT 1194
1071 CTGCGG-----CTCACGTGATGTCCTTACCAAGACCATGCTTAAAGTTAAAGCTT 1124
1195 CTGCGGCGATGCGCGCGGCGGAGTCCCGACCAAGAGACCACTCCCAAGCTCAAGCGCT 1254
1125 GAGTATGATCTTGAAGACAGGCTTTAAGTTGTATCCACCAATGATGATGATGATG 1184
1255 CGGATGATCTTCAACGAGACGCTGCGCTGTACCGCGCGGCGGCGGCGGCGGCGG 1314
1185 CGCTAATCGATGATGAGCTAG-----GAGGTTACAAAT 1220
1315 CGCAAGTTGACGCTACCTTCCGCGGCGGCGGAGCGGAGCGGAGGAGGATCATAT 1374
1221 CGCATGAGCAAGGAGCTTCAATCCCAATGAGGCTGATGATGATGATGATGATGATG 1280
1375 CCGGCGGACAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1434
1281 GGGTATGAGCTGAGACGATTTCAATCCAGCTCGTTTGGATGAGTCCCGCTGCTGC 1340
1435 GGGGCGGACCGCGGCGGCGGAGTTCACCGGAGAGTTCGCGAGCGCGCGGCGGCGG 1494
1341 CAAGACCGCGTGGCTTCAATCCCAATGAGGCTGATGATGATGATGATGATGATG 1400
1495 GAGAGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1554
1401 TCTTGTATGATGATGAGGCAATTTGACATCGCTGATGATGATGATGATGATGATG 1460
1555 CCGTCCGATCTTCAAGGCGAAGGCTCAACATGAGGCGGCTCTCTCAAGCGCTGAGC 1614
1461 CTGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1520
1615 GCTCTGCGGACCTTACGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1674
1521 TGACCAATCACTTCCG 1539
1675 CCGCGCGTGTGATCTTCCG 1693

RESULT 4
AX467615
LOCUS AX467615 1539 bp DNA linear PAT 16-UTL-2002
DEFINITION Sequence 3 from Patent WO246377.
ACCESSION AX467615
VERSION AX467615.1 GI:21900803
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosoids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 Kirochko, J.E., Cutler, A.J. and Abrams, S.R.

TITLE Abscisic acid 8'- and 7'- hydroxylase genes and related sequences
 JOURNAL Patent: WO 0246377-A 3 13-JUN-2002;
 National Research Council of Canada (CA)
 FEATURES Location/Qualifiers
 source 1.1539
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"
 /note="var. Columbia"
 ORIGIN
 Query Match 20.4%; Score 318.8; DB 6; Length 1539;
 Best Local Similarity 52.8%; Pred. No. 3,7e-80;
 Matches 761; Conservative 0; Mismatches 672; Indels 9; Gaps 3;
 QY 90 GTTATGCTGAGACCAAGAAAGATTGAAACATTTCTTAACAGGAAATTCAGAGTCC 149
 DB 78 GGTCTGATTAACCAAGATGCTTGAAGTCTTTGAAAGACAGGGTCTTACCGGAC 137
 QY 150 TCCCTATCATTTCTTCATCGAATGTTAAAGAACTTTGGAAATGATGCTTAAGCTTC 209
 DB 138 TCCCTAACCTCTCTGCTGAGATATAAGGAAATGTTGATGATGATGAGAGAG 197
 QY 210 TTCTCATCTATGCTTCTCTCACATATTTCTTCTAGAGTTCTCTCTTTTACCATCA 269
 DB 198 ATCTAAACCATTAATGTAACGATGATATACCCCAAGTCTCTCTCTGCTTAA 257
 QY 270 CTGAGAGAAATCTACGGTGTACATTTCTGTTGTTGGTCCGCAATTTCCGTTAC 329
 DB 258 GATGCTCATTTCTACGGAAGACATTTCTTCAATGATGAGACCATCTCCAGATGT 317
 QY 330 GGTAGCCCATCTCTGATTTATGAGAGATCTTCTCTAGTCTGAGTTCTACGAGAGAA 389
 DB 318 GATTAAGAAATCTGACAGATCAAGAGATCTTAAATAGTCAACGACCTTGGAGAAAC 377
 QY 390 TGAAGCTCAACCTTTGTTAAACAATTGAAAGCGATGACTTAAGTCTCAAGGTGA 449
 DB 378 TTCTCATTTCTTGTATGATGATGTTGATGAGG---TGGGCTTGAAGTTTCAAGGAGA 434
 QY 450 AAAATGGCTCATCATGAAAAATCTATGCCCCCTTCTTCAATGAGAAATCTTAAGT 509
 DB 435 TAAATGGGCGATCAAGAGAGATCAACCCGCTTTTCACTCGAAAAATCAAGAA 494
 QY 510 GCTTGAACAGTGTGTGAAGGTGATGATGATGATGATGATGATGATGATGATGATG 569
 DB 495 CATGATCCCTGCTTCAACATTTGTGACAGAGGTGTCTGTCATGAGAGAAAGCTATT 554
 QY 570 ATCAGAAACG---GTGAAGTTGAGGTAGATGTTGATGAGTGTTCAGATTTTGACTGA 626
 DB 555 TACAGATTAAGAAATGCGCTCTTGAAGTCAATGTTGGGCTTGTGAATTAAGTACG 614
 QY 627 AGATGTTATTAAGTAAACAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 686
 DB 615 GGAATCATCTCATCTGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 674
 QY 687 ACTTCAAGTCAACAAATGCTTTTGTGCTGAAGCTTTTCAAAAAGTCTTCAATCTCTGG 746
 DB 675 ACTTCAAGGGAATTTGGCTGAGCTTATGCAACAGCTTTTAAAGAAATCTTAACTCC 734
 QY 747 CTATAGATTTTTCGCAAGAGGAAATTTGAAGTCTCGAAGTGAAGCAAGAGATTAAG 806
 DB 735 ATCAGAGTTTTCACCAAGAGCAATAGAGATGAAGCAATAGATGAAGAGATGA 794
 QY 807 GAAATGTTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 866
 DB 795 CGTAAATTAAGAGGTGTTGAGCAAAAGGAGAAAGGAGAGAGAGTGAAGTGAAGTGA 854
 QY 867 AGAATGTAAGAGCCGCGCGGAGAGATTTGTTGGATTAATGA---TTCAAGCAAGAA 923
 DB 855 AAACGATGATTTGTTGGGATATTTGCTTGAATCAATTCAGAGAAATTCAGAGAAATGG 914
 QY 924 TGTACGGTTGAGACATTTGAGAGAGTGAAGAGCTTTTCTTCTGCGCGGAGAAACAGAC 983

DB 915 AATGAGCTGAGAGATGATGAAAGATGCAAGTGTGTTTATTTGCGCGCAACAGAC 974
 QY 984 AACTTTAATCTGCTGAGCGTGAAGACCATCTTCTATCATCATCACCCGAGTGGAGGC 1043
 DB 975 AACTTCAGTACTTTTGTGTGTGAGCTATGTTTATTAAGCATCACCAAGACTGGCAAC 1034
 QY 1044 CAAGCACCGTATAGAGTCCCTGAGGGCTCGAGCTCAAGTATGCTTCAACAGAGCA 1103
 DB 1035 TCGGCGACGAGAGAGATGATGAGTACCTCGGTGAAGAAATTAAGCTGATGAGTTC 1094
 QY 1104 TGTGTTAAGCTTAAACGTTGATGATCTTGAACGAGTCTTTAAGTGTATCTCAC 1163
 DB 1095 CTTTACACACTTAAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1154
 QY 1164 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1223
 DB 1155 GGTGCTGAGCTTAAAGATGTTGCAACAAAGAAATGATGATGATGATGATGATGATG 1214
 QY 1224 ATGTGCAAGAGCTTCTAATCCCATCATAGCGTCCATCATGACCAAGCATTTGGGG 1283
 DB 1215 AGCTGATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1274
 QY 1284 TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1343
 DB 1275 CGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1334
 QY 1344 AACCCCGTGGCTTCAATCCCATCATAGCGTCCATCATGACCAAGCATTTGGGG 1403
 DB 1335 GAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1394
 QY 1404 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1463
 DB 1395 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1454
 QY 1464 GGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1523
 DB 1455 CTCTCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1514
 QY 1524 AC 1525
 DB 1515 TC 1516
 RESULT 5
 BT003388 1614 bp mRNA linear PLN 29-JAN-2003
 LOCUS Arabidopsis thaliana cytochrome P450 (At3g14610) mRNA, complete cds.
 DEFINITION
 ACCESSION BT003388
 VERSION BT003388.1 GI:28059361
 KEYWORDS FLI_cDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 REFERENCE
 AUTHORS
 SCUTWICK, A., NGUYEN, M., TRIPP, M., PALM, C. J., JONES, T., WU, T.,
 CARMICHAEL, P., CHEN, H., CHEN, R., CHAN, M. M., CHANG, C. H., DALE, J. M.,
 DENG, J. M., HAYASHIZAKI, Y., HSUAN, V. W., LEE, J. M., ISHIDA, J.,
 KAMIYA, A., KAWAI, J., KIM, C. J., NATSUKAWA, M., ONODERA, C. S.,
 QUACH, H. T., SAKURAI, T., SATOH, M., SEKI, M., SHIM, P., TANG, C. C.,
 TORUMI, M., WONG, C., WU, H. C., YAMADA, K., YU, G., YUAN, S.,
 SHINZAKI, K., ECKER, J., THEOLOGIS, A. and DAVIS, R. W.
 Direct Submission
 Submitted (29-JAN-2003) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 COMMENT
 e-mail for correspondence: arab@sequence.stanford.edu
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RFL cDNAs (RFL cDNA: 'RIKEN

Arabidopsis Full-length cDNA: Seki, M., Narusaka, M., Ishida, J., Satoh, M., Kamliya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salik, Stanford, PGBC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Nguyen, M., Tripp, M., Southwick, A., Palm, C. J., Jones, T., Wu, T., Chen, H., Cheuk, R., Chan, M. M., Chang, C. H., Dale, J. M., Deng, J. M., Heuan, Y. N., Lee, J. M., Kim, C. J., Quach, H. L., Odeh, C. S., Shim, P., Tang, C. C., Theoumi, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Yuan, S., Ecker, J., Thoulas, A. and Davis, R. W.

Nguyen, M., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R. W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

source

1. 1614

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="U15533"

/ecotype="Columbia"

/note="This clone is in pENTR/SD-topo. This is a cloned PCR product using RIKEN clone RAVL04-16-M02 (AY072303) as a template"

gene

1. 1614

/gene="At3g14610"

/note="synonym: MIE1.11"

CDS

1. 1539

/gene="At3g14610"

/codon_start=1

/product="cytochrome P450"

/protein_id="AA030051.1"

/db_xref="GI:28059362"

/translation="MSFSVVAALPLVAVVLTMRIVKWKIKEMLESSLKQGLT
GTPYPLVGDIKRVNMMKARSPINVDIDITPLRLPALMLSHGKTFEIIWGL
PTIITNPDIKEVFNKVNDEKASTPELIRLLAGLASYGDKWASRRILINAPHL
EKIKNMI PARYHCCSEYVCOMKELPTDKESPLVYVPMVYNMADIVSHARSSYK
EGGRIFOLGELALBLIQAQKSYIPGSPYPTKSRKMAIDREVDVILGIVSKK
KAREGAPANDDLILILSNSESQGNMSVEDWKRCKLYFEAGDSTVLLVMT
VLNHHODQAFARAEVYQVLGENNPDDESJLNLKVMIMFNEVLRVLPVADQKVF
VNMKMKGLTLPAGIOIYLPILIVORDLEWGDAAKPERPFDGSKATKQVDF
FPGMGPRICIGQNFALFKXAMALILCKSFELSPSYHAPQVMTTRPQFAHLI
LHKL"

ORIGIN

Query Match 20.4%; Score 318.8; DB 8; Length 1614;

Best Local Similarity 52.8%; Pred. No. 3.7e-80;

Matches 761; Conservative 0; Mismatches 672; Indels 9; Gaps 3;

90 GTTATGTGAGACCAAGAAAGATTGAACAATTCCTTAACAAGAAATCGAGTCC 149
78 GGTCTGGAATAAACAAGATGCTTGAAGTCTTTGAAAAGAGAGGCTTACAGGAC 137
150 TCTTATCATTTTTCATCGGAATGTTAAGACTTTTGGATGATGCTTAAGCTTC 209
138 TCTTGAACCTCTCTCGGAGATTAAGAGAAATGTTGATGATGATGATGAGCGAG 197
210 TTCTCATCTCATGCTTCTCTCAATATCTTCCAGAGTCTCTCTTTTACATCA 269
138 ATCTAAACCATCAATGATGAAGATGATATCAACCCAGCTCTCTCTCTGCTTAA 257
270 CTGAGAAAATCTACGCTGCTACATTTCTGTTGGTTCGCTCAACTTCGGGTAC 329
258 GATGCTCATCTTCAACGAAAGACTTCTGATGATGATGATGATGATGATGATG 317
330 GGTAGCGATCTCTGATTTGATCAAGAGATCTCTCAAGTCTGATTTCAAGAGAA 389
318 GATTAACGATCTCTGAGAGATCAAGAAAGCTTTTAATTAAGTCAACGACTTTGA 377
390 TGAAGCTCACTCTTGGTTAAACAATTGAAGCGATGATCACTTAAGTCAAGAGTGA 449

378 TTCTACATTCCTTTGATCAGATTGTTAGCAG--TGCGCTTGCAAGTTTCAAGGAGA 434
450 AAAATGGGCTCATCATCAAAAATCATTAAGCCCTACTTTTCATATGAGAGATTAGTT 509
435 TAAATGGGAGATCAGAGAGATCATCAACCGCGCTTTTCACTCGAAAAATCAAGAA 494
510 GCTTGATCAAGTGTGTGTAAGAGTGTGATGATGATGATGATGATGATGATGATG 569
495 CATGATCCCTGGCTTCAACATTTTGAGAGAGATGTTGTCTCAATGGAGAGATTT 554
570 ATCAAGAAAACG--GTGAGTTGAGTAGATGCTATGATGATGATGATGATGATGATG 626
555 TACAGATTAAGAAATGCGCTTGAAGTCAATGTTGGCTTGGCTTGAATATGATGCTG 614
627 AATGTTATTAGTAACAAGCTTTTGAAGTATGATGATGATGATGATGATGATGATG 686
615 GATGTCATCTCACTACTGCTTTTGAAGTATGATGATGATGATGATGATGATGATG 674
687 ACTTCAAGCTCAACAAATGCTTCTTGTGCTGAAGCTTTCAAAAAGTCTTCAATCTCG 746
675 ACTTCAAGGGGAATGGCTGAGCTTATCGCAAGCTTTTAAGAAATCTTACATCTCG 734
747 CTATGATTTTTCGCAAGAGAGAAATTTGAAGTCTCGAAGTTAGCAAGAGATAG 806
735 ATCGAGGTTTACCAACAAAGAGCAATAGAGATGAAAGCAATGATGAGAGATGA 794
807 GAATGCTGTTGTAAGCTGATAGAGCGCGAGACAAACGCTATGATGAGAGAGGGA 866
795 CGTAAATTTGAGAGGATATGAGCAACCGGAGAAAGCGAGAGATGAGAGACCAAG 854
867 AATAATGAGAGAGCGCGGAGAGATTTGTTGGATTAATA--TTCAAGCAAGAA 923
855 AATGATATTTGTTGGGATATTTGCTTGAATCAATTCAGAGAAATCTCAAGAAATGG 914
924 TGTGACGTTGAGCAATTTGAGAGAGTGTAAAGCTTTTCTTCCCGGAGAAACAGAC 983
915 AATGAGCGTAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 974
984 AACTTCTAATGCTGAGAGTGAAGACACATCTTGCATACAGACCGGAGTGGAGGC 1043
975 AACTTCAATCTTTTGTGTGAGTATGATGATGATGATGATGATGATGATGATGATG 1034
1044 CAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1103
1035 TCGCGCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1094
1104 TGTGCTTAAGCTTAAAGCTTGAATGATGATGATGATGATGATGATGATGATGATG 1163
1095 CTTTAAACCTTAAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1154
1164 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1223
1155 GGTGCTGAGCTTAAAGCTTGAATGATGATGATGATGATGATGATGATGATGATG 1214
1224 AATGCGACGAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1283
1215 AACTGAAATTCAAATTTACTTACCACTTATCTTGTTCAGCTGACACCGAGCTTGG 1274
1284 TAATGACGTGAAGAAATTCACAGCTGCTGCTTGGAGTGAAGTGAAGTGCAGTGC 1343
1275 CGATGATGACAGGAGATTTAAACCGAGCGGTTCAGAGAGGCTCTCAAGAGCAACA 1334
1344 ACACCGGCTGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1403
1335 GAACGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1394
1404 TGTATACTTACGACCAATTTGACCTGCTGATGATGATGATGATGATGATGATGATG 1463
1395 TGTATGTTGAGAGCAAGATGAGCAATGAGCTTGAATTCACAAAGCTTCTTCTGAG 1454
1464 GGTCTCTTATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1523
1455 CTCTCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1514

QY 1524 AC 1525
 Db 1515 TC 1516
 RESULT 6
 AY072303
 LOCUS
 DEFINITION Arabidopsis thaliana cytochrome P450 (At3g14610; MIE1.11) mRNA,
 complete cds.
 ACCESSION AY072303.1 GI:18252154
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1786)
 Nguyen, M., Karlin-Neumann, G., Southwick, A., Lam, B., Miranda, M.,
 Palm, C. J., Bowser, L., Jones, T., Barn, J., Carninci, P., Chen, H.,
 Cheuk, R., Chung, M. K., Hayashizaki, Y., Ishida, J., Kawai, J.,
 Kawai, J., Kim, C., Lin, J., Liu, S. X., Natsumura, P., Pham, P. K.,
 Sakano, H., Sakurai, T., Saitou, M., Seki, M., Shim, P., Yamada, K.,
 Shinzaki, K., Ecker, J., Theologis, A. and Davis, R. W.
 Direct Submission
 Submitted (02-JAN-2002) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 e-mail for correspondence: arab@sequence.stanford.edu
 COMMENT
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
 Arabidopsis Full-length cDNA'): Seki, M., Natsumura, M., Ishida, J.,
 Saitou, M., Kawai, J., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinzaki, K.
 The Saik, Stanford, PGSC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Nguyen, M.,
 Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M.,
 Bowser, L., Jones, T., Barn, J., Chen, H., Cheuk, R., Chung, M. K.,
 Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shim, P.,
 Yamada, K., Ecker, J., Theologis, A. and Davis, R. W.
 Nguyen, M. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed
 equally to this work. Shinzaki, K. (RIKEN GSC) and Davis, R. W.
 (SSP/Stanford) contributed equally to this work as PIs.
 FEATURES
 source
 1. 1786
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="RAFL04-16-M02"
 /ecotype="Columbia"
 /note="This clone is in pBluescript"
 1. 1786
 /gene="AC3g14610; MIE1.11"
 100. 1638
 /gene="AC3g14610; MIE1.11"
 /product="cytochrome P450"
 /codon_start=1
 /protein_id="AA161910.1"
 /db_xref="GI:18252154"
 /translation="MSFVVALPVLVAVVLTWIRIVKMWIKRKLKSLKRGIT
 GTPYPLVGDIKAVDMMEARKSPINTVDITPEIRLPLAKMLNSGKFFIIGL
 PIVITNPEQIKVENKVNDFKASTPILRLAGGLASVGDMAKSHRIINAFHL
 EKIKNMIPAFHYCSEVUCOMEKFTDESPLVAVMLVMTADYSHAFSSYK
 EGORTFQDGEIABELIAQAFKSYIPGSRFPYTKSNREMKLIDSEVVLIGIVSKRE
 KAREGEPPANDDLIGLLESSESGGMSYEDVMCECKFYRAGQEVLLIWMW
 VLSHQDQARAREVWQVLEENKPMDESLNKLKVTIMLFENYRLIYPVADLKV
 VKEMELGELTIPAGIYIYLPILVORDELMGDDAADFKEFRFDGSKATKQVSEF

ORIGIN
 LHLK"
 FPFMGPRICIGONFAMLEAKMALLIOKFSHELSPSYVHAPQVWTRPQGAHLI
 Query Match 20.4%; Score 318.8; DB 8; Length 1786;
 Best Local Similarity 52.8%; Pred. No. 3.8e-80;
 Matches 761; Conservative 0; Mismatches 672; Indels 9; Gaps 3;
 QY 90 GTATGGTGGAGACCAAGAAAGATGAAGAACTTCTTAACAAGAAATCGAGTCC 149
 Db 177 GGTCTGATTAACCAAGATGCTTGAGACTTCTTGAAAGACAGGCTTACCGGAAC 236
 QY 150 TCCTTATCATTTCTTCATCGGAAATGTTAAGAACTTTGGATGATGCTTAACCTTC 209
 Db 237 TCCTTACACTCTCTCGTCCGAGATTAAGAGAAATGTTGATGATGATGAGAACCGAG 296
 QY 210 TTCTCATCCTATGCTTCTCTCTCAATATCTTCCAGAGCTCTCTTTTACATCA 269
 Db 237 ATCTAAACCATCAATGATGAACGATGATACACCCACGCTCTCTCTCTTCCCTTAA 356
 QY 270 CTGAGAAATCTACGATGCTACATTTCTGTTGGTTGGTCCAACTTCCGTTAAC 329
 Db 357 GATGCTCAATCTCACGGAAGAAAGACTTCTCATATGATGAGACCACTTCCAAAGATTGT 416
 QY 330 GGTAGCCGATCTGATTTGATCAGAGAGATCTTCTTAAGTCTGAGTTCTAGAGAA 389
 Db 417 GATACGAAATCTTACGAGATCAAGAGAACTTTAAATGAATCAAGACATTTAGAAAGC 476
 QY 390 TGAAGTCAACCTTGGTTAAACAACCTGAAGCGATGAGACTTATGCTCAAGGTGA 449
 Db 477 TTCTACATCTCTTGTGATGATGATGATGAGAG---TGGCTTGCAAGTTACAGGAGA 533
 QY 450 AAAATGGCTCATCATGAAATATCATTAAGCTTCTTCAATGAGAACTTAAGT 509
 Db 534 TAAATGGCGAGTACAGAGAGATCAACCCGCTTTCACCTGCAAAAATCAAGAA 593
 QY 510 GCTTGTACCACTTGGTTGAAAGATGAGTATGATGATGATGATGATGATGATGAT 559
 Db 554 CATGATCCCTGCTTCTACATTTTGGAGAGAGTGTCTGTCATAGGAGAAAGCTATT 653
 QY 570 ATCAGAAACG---GTGAAGTGAAGTATGATGATGATGATGATGATGATGATGAT 626
 Db 654 TACAGATTAAGATCGCTCTTGAAGTGAAGTGTGGCTTGTGATGATGATGATGAT 713
 QY 627 AGATGTTATGATGAAGAGCTTTTGAAGTATGATGATGATGATGATGATGATGATGAT 686
 Db 714 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773
 QY 687 ACTTCAGCTCAACAAATGCTTCTTGTGTAAGCTTTCAAAAAGCTTTCATCTCTG 746
 Db 774 ACTTACAGGGAATGCTGAGCTTATGAGCAAGCTTTTAAAGAACTTACATCTCTG 833
 QY 747 CTATGATTTTTCGCAAGAGGAATTTGAATCTTCGGAAGTTTACAGAGATTAAG 806
 Db 834 ATCGAGGTTTACCAAGAGAGCAATGAAGATGAAGCAATGAATGAAGATGAAG 893
 QY 807 GAAGTCGTTGTAAGTGAAGTGAAGCGCGAGACAAACGCTTATGATGAGAGAGGGA 866
 Db 894 CGTATATTTGAGAGTATTTGAGAGAAACGGAGAGAGAGAGAGAGAGAGAGAGAG 953
 QY 867 AGAATGTAAGAGCGCGAGAGAGATTTGTTGGAATTAAGA---TTACGAGAAAGAA 923
 Db 954 AAACGATGATTTGTTGGAGATATTTGCTTGAATTCATTCAGAGAAATCTCAAGAAATGG 1013
 QY 924 TGTGACGTTTCAGACATTTGTAAGAGATGAAGAGCTTTTCTTCCGCGGAAACAGC 983
 Db 1014 AATGACCGTGAAGATGATGAAGAGATGAAGAGTGTATTTATTTCCGGGCAAGAGAC 1073
 QY 984 AACTTCAATGCTGAGTGAAGCAACATCTTGTATTCATGACACCGGAGTGGCGAGC 1043
 Db 1074 AACTTCAATGCTTGTGAGTGAAGCAACATCTTGTATTTAAGCAATCAACAGCTGGCAGC 1133
 QY 1044 CAAAGACGATGAGTCTCAGAGGTCTGCGGCTCAGCTGATGTCTTCAAGAGAGCA 1103

Db	1134	TCGGGACAGAGAGAGATGATCAAGTACTCGGTGAAAATATATAAAGCTATATGAGTCC	1193
Qy	1104	TGTCTTAAGCTTAAAGCTTGAAGTATCTTGAAGAGCTTAAAGGTTGATCCACC	1163
Db	1194	CCTTAACAACCTTAAAGTATGATATATCTTCAATGAGGTTTGGGCTTAACCTCC	1253
Qy	1164	AATAGTACGATGATTCGACGGCTTAAATCGATGTGAGAGGAGGTAACAATGCC	1223
Db	1254	GGTGGCTCAGCTTAAACAGAGTGTGCAAAAGAAATGAGAGCTCGAAGAGTTGACCTCC	1313
Qy	1224	ATGTGGCAGGAGCTTCTATATCCCATCTAGCGGCTCATATATACCAAGCAATTTGGG	1283
Db	1314	AGCTGGAATTCAAATTTACTTACCAACTATCTTGTCCAGCGTGAACCGAGCTTTGGGG	1373
Qy	1284	TATATACGTGAAGCAATTCATTCAGCTCGGTTTGCAGATGAGTGGCGGCTGCTGCCAA	1343
Db	1374	CGATGATGACGCGATTTTAAACCGAGCGGTTGAGACGCGGCTTCAAGGCAACAA	1433
Qy	1344	ACACCCCGTGGCTTCAACCGTTGGCGCTCGAGTTCGTACATGATTCGTCAATCT	1403
Db	1434	GAACCAAGTCTCTCTTCTCCCTTGGATGGGAGCTGAGATTTGCACTGGTCAAGAAATT	1493
Qy	1404	TGCTATCTTACGCGCAATGACACTGCTGTATATGATGACAGCTTCACTTCACTT	1463
Db	1494	TGCTATGTTGAGGAGCAAGATGAGCAATGAGCTTTGATTTCTACAAAGTTCTCTCGAGCT	1553
Qy	1464	GGCTCTACTTATACAGCATGACCTACCGTCTCTTATGTTGCTTATCTCTCAACATGGTC	1523
Db	1554	CTCTCTCTTATGTTTACAGCGCGCTCAACAGTCAATGACCACTGCTCCCAATTCGAGGC	1613
Qy	1524	AC 1525	
Db	1614	TC 1615	
RESULT 7			
AF446870		1548 bp mRNA 1linear	PLN 06-DEC-2001
LOCUS	AF446870	Arabidopsis thaliana AT3g14620/MIR1_12 mRNA, complete cds.	
DEFINITION	AF446870		
ACCESSION	AF446870.1	GI:17386113	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			

COMMENT
<p>RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'). Seki, M., Narusaka, M., Ishida, J., Saitou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.</p> <p>The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Meyers, M.C., Shinn, P., Banh, J., Bower, L., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.</p> <p>Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.</p>
<p>Location/Qualifiers</p> <p>1..1548</p> <p>/organism="Arabidopsis thaliana"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:3702"</p> <p>/chromosome="3"</p> <p>/clone="U13574"</p> <p>/ecotype="Columbia"</p> <p>/note="This clone is in pUNI 51"</p> <p>1..1548</p> <p>/note="putative cytochrome P450"</p> <p>/codon_start=1</p> <p>/product="AT3g14620/MIR1_12"</p> <p>/protein_id="AAL38603.1"</p> <p>/db_xref="GI:17386114"</p> <p>/translation="MSDRTKISAVVAVVVTVVWIKGNLVNMLRPKNKAYLK...ROGSGTPTFLVGDIKREASMEVQESRPIPLVDYTHFVMPLIQTVADHDKTSYM...WGPASIVTIVPHEIKDYNRVDPKPPVPELVEFATGVALYEGEKMSKRIIN...PSFHEKIKMIPAFYSCSEMSKMKIKITGEOSSNEIDVMPYGLDTSVSRXF...GSSYBGRITFELBOGRVITKLEAFIPQGRFLPRTKNLIRWQINKEVSLRREI...IMKRGKNDTGEAPKNDLIGLLESNSGDHMSJEDVBCRLPHFPAQCTTPTLLW...TMIMSHKQMDQAPREILKIVIGIKNNKPNFDALSRKTSMTLNEVLRAPPEITLG...RVMEKRTIGEDMLPGGAQVIVIVLWVRDPELWGEDVHFNERNFADDISATKNO...VSPFLPWGPRFCPOGNPALMEAKVALVILQRPSELSPTHTAPHTVLHLPOGA...PIIFHML"</p>
<p>ORIGIN</p> <p>Query Match 18.7%; Score 291.6; DB 8; Length 1548;</p> <p>Best Local Similarity 52.2%; Pred. No. 2.5e-72;</p> <p>Matches 770; Conservative 0; Mismatches 689; Indels 15; Gaps 5;</p>
<p>70 ATATGAGAGGATGCTCTGTATGAGGAGACCAAGAAATGGAAGACATTTCTCT 129</p> <p>73 ATATGAAAGGCTAATGATGCTTGGCTTACCAAGAAATGAGCTTATCTAAG 132</p> <p>130 AAACAAGATTCAGAGTCTCTTATCATTTCTTCATGCGAATGTTAAAGAACTGTT 189</p> <p>133 AGACAAGGCTCTCGAAGCTCTTCACTTCTTGGGAGACATTAAGGAAGCC 192</p> <p>190 GGAATATGCTTAAAGCTTCTTCATCTCAATGCTTCTCTCAATATTTCTTGA 249</p> <p>193 AGTATGATGAGCAAAAGTCAAGACCAATCATCTAATGATGATTAACGATGCT 252</p> <p>250 GTTCTCTCTTTTAAACATCATCTGAGAAATCTACCGTGTCACTTTCTGTTTC 309</p> <p>253 GTATGCTCTGTATCAACAAGACCGTTAAGATCATGGAAGACATCTTAACATG 312</p> <p>310 GGTTCACATTTCCGTTACCGTATGCGATCTGTTGATGAGAGAGATCTTCTAAG 369</p> <p>313 GGAACATATGAGAGTGTATGAGCAACACATTAAGATGTTTAAACGA 372</p> <p>370 TCTGAGTTTACAGAGATGAGATGATCAACCTTTGTTAAACAATTGAGGATGGA 429</p> <p>373 GTTACGATTTCCGAAAGCGGAGTGAACCAATCGTT--GAGCTTCCGAGCTGA 429</p>

QY	430	CTACTTAGTCTCAAAAGGTGAAAAATGGGGCTCAGCATCATGAAATAATCAATTAGCCCTACCTTT	489
Db	430	GTGCGCGCTGTATGAGAGAGAAATGGTCTTAAGCAAGGAAGATTAAACCTTTCTTTT	489
QY	490	CATATGAGAGATCTTAAAGTTGCTTGTAACAGTTGTGTGGAAGGTGACTGATATGTGT	549
Db	490	CATTGTGAAAAGCTCAAGATTATGATACCTGACATTCTACAGAGAGCTGACGGAGATGATA	549
QY	550	GATTAATAGTCCGATAAGTTATCAAGAAAACGTTGAA---GTGAGAGATATGCTATAGAG	606
Db	550	AGTAAATGGGAGAAAGTTGGTGACCGACCAAGGATCATCGAATGATCGATGTTGGCCA	609
QY	607	TGGTTTCAGATTTTTCAGTCAAGATGTTATTAGTAAACAGCTTTTGGAACTACTATGAA	666
Db	610	TATCTTGGAGATTTAACATCGATGTAATTTCAAGTACTCGATTTGGTATCTACTACGAG	669
QY	657	GATGTCGAGCAGTTTTCGACTTCAAGCTCAACAAATGCTTTCTTTGTGCTGAAGCTTTT	726
Db	670	GAAAGTAAAGAAATCTTTGAGCTTCAAGAGAAACAAAGCCGGCCAGTGTAAAGCTCTT	729
QY	727	CAAAAAGCTTCATTTCCCGCTATATAGATTTTTCGCAACAAGGGAAATTTGAAGTCCGG	786
Db	730	GAGTTGGCTTTATCTCCCTGGAATGAGGTTTCTACAAACAAAGCAATTTGAGATGAGA	789
QY	787	AAGTTAGACAAGAGATPAGAGAGTCGTTGTGTAAGCTGATPAGACCGGCGAGACAAAC	846
Db	790	CAAAATTAACAAGA---AGTGAATCTTGAAGTGAAGAGATCATGAAAGAAGAGAG	846
QY	847	GCTATAGATGAGAAAGGGGAAGATGTPAAGAGCCGGCCGCGAGAGATTTGTGGATTA	906
Db	847	GGTATAGATTAAGGGGAGGCTCCAAAGATATCTGTTPAAGAA--TACTTTGGAAATCG	903
QY	907	ATGATTCAGGCCAAAGAAATGTGACGGTTACAGACATTTGAGAGAGTGTAAAGCTTTTC	966
Db	904	AATTTCGAGATCATGATATGATATGAAATATGATGTAAGAAAGTGAAGGCTGTTTCAC	963
QY	967	TTGCGCGGGAAAACAGACAACTTCTAATCTGCTGACGTGAGACACATCTTGCTATCATG	1026
Db	964	TTGCGGGGGCCAAAGAACACCGCGGTGCTGCTCGTGTGAGACTATGATATATGTAAGCAT	1023
QY	1027	CACCGGAGTGGCAGAGCCAAAGACAGTATGAGGCTCAGAGTCTGCGGCTCACATGAT	1086
Db	1024	CACCAAAATGTGCAAGATCAAGCTTGAAGAAAGATATTAAGTCAATTTGGSAAGAACAC	1083
QY	1087	GTCCCTACCAAGACCATGTGCTTAAGCTTAAACGTTGAGTATGATCTTGAACGATCT	1146
Db	1084	AAACCTAATTTTGAAGCACTCTCGACTCAAAACAAATGACATGATCTTGAAGAAATT	1143
QY	1147	TTAAGTTGATTCACCAATAGTACTACGATTTGACGGCGCTTAATCGATGTGAAGCTA	1206
Db	1144	TTGAGGTTATATCCACCGAGATCTACTTGTCGCAACCTTATGAGAAAGAAACAAAGCTT	1203
QY	1207	GGAGGGTAC---AAATCCCATGTGSCACGAGCTTCTAATCCCAATCATAGCGGTCCAT	1263
Db	1204	GGTGAAGATGACACTACAGAGTGGTGCAAGATGATGATTTCCGTGTTCTATAGGTTGAT	1263
QY	1264	CATGACCAAGCCATTTGGGGTATATGACGTGAACGAATTCATCCAGCTCGGTTGGCGAT	1323
Db	1264	CGTGAATCCGAGCTCTGGGGTGAAGATGTGCAATTCATTCGAAAGATTCGCAAT	1323
QY	1324	GGAGTCCCGGTGTGCCCAACACCCCGTTGCTTCATACCGTTTGGCTCGAGTTGCT	1383
Db	1324	GGTATTCAAAAGCCCAAAAGAACAAAGTTTGTTCTTCGTTTGGATGGGACCAAGA	1383
QY	1384	ACATGCAATGTGCAGAACTTGTCTATCTTCAGGCCAAATTGACACTCGCTGTATGATC	1443
Db	1384	TTCTGTCTGTGGCCAGAAATTTGTTGAATGGAAGCTTAAGATGACCTGTGTTGATTTG	1443
QY	1444	CAACGCTTACCTTCACTTGGCTCTATCTTATCTTACAGATGACACTACCGCTTATGTTG	1503
Db	1444	CAGAGATTTCTCTTTAGCTGTCTCTTCTTGTCATCGACGACCTCATACGGTCTTTACA	1503
QY	1504	CTTTATCTCTCAACATGATGACCAATCACCTTCC	1537

Db 1504 CTTGACCTCAGTTTGAGTGCTCCGTTGATCTTCC 1537

RESULT 8	AY052208	1784 bp	mRNA	linear	PLN 09-SEP-2001
LOCUS	AY052208				
DEFINITION	Arabidopsis thaliana AT5g14620/MIE1_12 mRNA, complete cds.				
ACCESSION	AY052208				
VERSION	AY052208.1	GI:15529168			
KEYWORDS	FLI CDNA				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1 (bases 1 to 1784)				
AUTHORS	Chen,H., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Barth,J., Bowers,J., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pam,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.				
TITLE	Arabidopsis cDNA clones				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1784)				
AUTHORS	Chen,H., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Barth,J., Bowers,J., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pam,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-AUG-2001) Salk Institute Genomic Analysis Laboratory (SICGL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA				
COMMENT	US: RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN Arabidopsis Full-length cDNA) : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.				

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Shim,P., Bam,J., Bowers,J., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pam,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Chen,H. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers

1. .1784

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="RAFL06-77-B22 (R13574)"

/ecotype="Columbia"

1. .53

54. .1601

/note="putative cytochrome P450"

/codon_start=1

/product="AT5G14620/MIE1_12"

/protein_id="AAK97679.1"

/db_xref="GI:15529169"
 /translation="MSDTKISAVAAVAVVTVTVTWIKGLNVAMLRPKNEAYLK
 ROGSGTPEFLVDIKREASNVQEKSPINLTDYTRVAPLLIQVXKHGTSTM
 WMGPASVITVPEPHIDVLRNRYDFKPPHIVELFATGVALYEGHMKHRIIN
 GSFLKXKIMIPAYECSSEMIKWEKIVTEGSSNEIDWVYGLDLSVY SRFA
 GSYEEGRKIFELQEOGRVLEKALFLAFPGMRPLPTKNLIRMOIKVEYSLREI
 IMKROMDGGEAKNDLIGILISNSGDMGSIIEVVEECILFHACQETVALIVR
 TMILSHQKQODARBEILKVI GKNNKNPDALSLKMTILNEVRLIPGILLG
 RTVEKRTLQEDMTLPGRQVVI PVMWEDBELMSEKEDMHEHNERFDGSKATIKQ
 VSLPFGMPKPCPGNFMALMEKMLVLLQKPSFELSPTTHAPRVLTHPQFA
 PLIFHML"
 1602..1784

3'UTR
 ORIGIN

Query Match 18.6% Score 290.6; DB 8; Length 1784;
 Best Local Similarity 52.0%; Pred. No. 5e-72;
 Matches 775; Conservative 0; Mismatches 699; Indels 15; Gaps 5;

70 ATAGTGAAGGATGCTCTGCTGTTATGCTGAGACCAAGAAAGTTGAAGACATTTCT 129
 126 ATATGGAAGAGCTGGAATGATGCTGCTTACCAAGAAAGTTGAAGACATTTAAAG 185
 130 AAACAGAAATTCAGAGCTCTCTTATCATTTCTTATGAGAAATGTTAAAGACTTGT 189
 186 AGACAAGGCTCTCTGGAACCTCTTACCTTTCTTGTGAGAACATTTAAAGGAGACC 245
 190 GGAATGATGCTTAAAGCTTCTTCTCATCTTCTGCTCTCTCAAAATATCTTCTAGA 249
 246 AGTATGCTGAGCAAGAAAGTCCAGACCATCAATCTAATGATGATTAACAGCATGT 305
 250 GTTCTCTTTTACCATCATGAGCAAAATCTAAGTGTCTACATTTCTGTTTGTTC 309
 306 GTCAATGCTCTGATPACACAGACCGTTAAGATCAAGAGACATTTACATGTGAGATG 365
 310 GGTCCAACTTCCGCTTACCGGTAGCCGATCTGATTTGATGAGAGATCTTCTTAAG 369
 366 GGAACCTATAGGAGGTGATGTCAGCAAAACCAAGAACATCAAGATGTTTAAACAGA 425
 370 TCTGATCTCTAGAGAAATGAAAGCTCAACCTTTGTTAAACATTTGAAGGACATGA 429
 426 GTTACGATTTCCGAGCCGCGCAGTGCACCAATGCT--GAGCTCTTCCGACATGA 482
 430 CTACTTACTTCAAAAGTGAATAATGGCTCATCATCAAAATCTTAGCCCTTACTT 489
 483 GTCCGCTGTATGAGAGGAGAAATGCTTAAAGCAAGAGATTTAAACCTCTTCTT 542
 490 CATATGAGGATCTTAAGTCTGTGTCAGTGTGTTGAAGTGTGCTGATATGCTG 549
 543 CATCTTGAAGGCTCAAGATTTATGATCTGCTTACAGAGGCTGCGAGCATGATA 602
 550 GATTAATGCTCCGATTAAGTTATCAGAAACGATGA--GTTAGGATAGATCTATGAG 606
 603 AGTAAATGAGGAAGTGTGCTGACCGACCAAGATCAATCAATGATGATGATTTGGCCA 652
 607 TGTGTTCAAGATTTGACGAGATGTTATAGTGAAGACGCTTTGGAAGTGTATGA 666
 663 TATCTTGAAGATTTAACTGATGATGATTTATTCAGTGTGATGATGATGATGATG 722
 667 GATGCTGAGGAGTTTTCGACTTCAAGCTCAACAAATGCTTTTGTGCTGAAGCTTT 726
 723 GAAGGTAGAGATCTTTGAGCTTCAAGAAAGCAAGCGCGGAGTGTATTAACCTCTT 782
 727 CAAGAAAGCTTCAATCCGCTGATAGATTTTTCAGCAAGAGGAAATTTGAAGTCTGG 786
 783 GATGCTGCTTTTATCCCTGAGATAGGTTTCTTACCAAGAAACATTTGAGATGAGA 842
 787 AAGTTAGCAAGAGATTAAGAGCTGTTTGAAGCTATAGAGCGCGAGACAAAC 846
 843 CAATTAACAAAG--AGTAAATCTAGACTGAGAGATCTATCTGAAAGACAGAG 899
 847 GGTATAGTGAAGAGAGAGATGTAAGAGCGCGCGAGAGATTTGTTGGATTA 906
 900 GGTATAGTGAAGAGAGAGCTTCAAGAAATGATCTGTTAGAA--TACTTTTGAAGATCG 956

QY 907 ATGATTCAGCAAGAAAGTGTACGCTTCAAGACATTTGTGAGAGATGTAAGACTTTTC 966
 DB 957 AATTCCTGAGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016
 QY 967 TTGCGCGGAGAAACAGCAACTTTAATCTGCTGAGCGAGACACATCTTGTATTCATG 1026
 DB 1017 TTGCGCGGAGAAACAGCAACTTTAATCTGCTGAGCGAGACACATCTTGTATTCATG 1076
 QY 1027 CAACCGGAGAGCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1086
 DB 1077 CACCAAAATGAGCGAGATCAAGCTAGAGAGAGATTTAAAGCTATTTGAGAGAGAGAG 1136
 QY 1087 GTCCTACCAAGAGACATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1146
 DB 1137 AAACCTTATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1196
 QY 1147 TTAAGTGTATTCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206
 DB 1197 TTGAGGTTATCCACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1256
 QY 1207 GGAGGATCA--AAATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1263
 DB 1257 GGTGAGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 1316
 QY 1264 CATGACCAAGCCATTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1323
 DB 1317 CGTGAATCCGAGCTCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1376
 QY 1324 GGAGTCCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383
 DB 1377 GGTATCTAAAGACCAAGAAACCAAGATTTGATGATGATGATGATGATGATGATGATGATGAT 1436
 QY 1384 ACATGAT 1443
 DB 1437 TTCTGCTGCGAGAAATTTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1496
 QY 1444 CAACGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1503
 DB 1497 CAGATATCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1556
 QY 1504 CTTTATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1552
 DB 1557 CTTCACTCTCACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1605

RESULT 9
 AK120018
 LOCUS
 DEFINITION
 Oryza sativa (japonica cultivar-group) cDNA clone:002-186-B11, full insert sequence.
 AK120018
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 1 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yabuchi, J., Ishikawa, M., Yamada, H., Ooka, K., Kojima, K., Nishikawa, T., Ohneda, E., Yabuchi, W., Suzuki, K., Li, C., Ohtsuki, K., Shihiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otsu, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kuroseki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikiura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Mura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,

TITLE	JOURNAL	REFERENCE
Science 301 (5631), 376-379 (2003)	22752273	12669764
Adachi, J., Aizawa, K., Akimura, T., Arahawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hoti, F., Hotte, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mitsuoka, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Narikawa, R., Nishikura, J., Nishii, K., Oka, H., Ose, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shizaki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akhihara, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.	Unpublished	3 (bases 1 to 1860)
Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kamondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)	Unpublished	3 (bases 1 to 1860)
URL: http://cdna01.dna.affrc.go.jp/cDNA/NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Negata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Oka, H., Hotte, I., Kojima, K., Namiki, T., Ohneda, E., Yanagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Yamamoto, M., and Nakahama, Y.	Unpublished	3 (bases 1 to 1860)
Full Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujiwara, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nishikura, J., Oka, M., Ryo, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tanoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.	Unpublished	3 (bases 1 to 1860)
Genome Exploration Research Group in Riken: Adachi, J., Aizawa, K., Akimura, T., Arahawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hoti, F., Hotte, I., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mitsuoka, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Narikawa, R., Nishikura, J., Nishii, K., Oka, H., Ose, N., Ota, Y., Otomo, Y., Ryo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shizaki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akhihara, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A., and Hayashizaki, Y.	Unpublished	3 (bases 1 to 1860)
Location/Qualifiers	1..1860	

ORIGIN											/organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /cultiVar="Nipponbare" /db_xref="taxon:39947" /clone="002-186-B11"
Query Match	18.6%	Score 290.2	DB 8	Length 1860							
Best Local Similarity	52.1%	Pred. No. 6.6e-72									
Matches 696	Conservative 0	Mismatches 633	Indels 6	Gaps 2							
QY	228	CTCTACCAATATTTCTCTGAGTCTCTCTCTTTTACCACTGAGGAAATCTACGG	287								
DB	282	CTGCACCGCATCGCGCCCGCGTCCCTTCTCTGAGACGCCGTACGACACACCG	341								
QY	288	TGCTACATTTTGGTTGGTTGGTCCCACTTCCGTTTACCGTACCGGACCTGATTT	347								
DB	342	TAACTGTCAATTCATTTGGTTGGTCCAGTCCAGCAAGATGATATTCGACCTGAAAT	401								
QY	348	GATCAGAGAGATCTCTCTAAGTCTGAGTCTCTAAGAGAAATGAACTCACCTTTGGT	407								
DB	402	AGTAAGAGAGGTTTTCATAGTTTGAACATTTGACAAACCAAGATGATCCGATGCG	461								
QY	408	TAAACATCTTGAAGCGCATGATCTTATCTTCAAAAGTGAAAAATGGGCTCATCATG	467								
DB	462	AAAGTTGCT--AGCGACAGAGATTGAGACTACGAAGCGAAGAAATGGGCTTAAGACCG	518								
QY	468	AAATATCATATGACCTTACTTTCAATATGAGAAATCTTAAGTTGCTTACCACTTGTGT	527								
DB	519	GAGATTTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	578								
QY	528	GAAAGTGTACTGATATATGTGTGATTAATGTGCTCCGATAAGTTATCAGAAAACGGTAACT	587								
DB	579	CAACTGTGACCGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG	638								
QY	588	TGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	647								
DB	639	TGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	698								
QY	648	TTTGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	707								
DB	699	ATTGTTGACGATTTCAAGAGGAGAGAAATTTTTCAGCTGACGAAGAGCTGGCTCA	758								
QY	708	TCCTTGTGCTAAGCTTTCAAAAAGCTTCATCTCTGCTATAGATTTTTCGACAAAG	767								
DB	759	ATATTTTACAGAGCCCTTCAGAAACTTTTATATACCAAGCTATGTGACTTCGCAACATA	818								
QY	768	AGGAAATTTGAAGTCTCGAAAGTTAGCAAGAGATTAAGAAAGTGGTTGTAAGCTGAT	827								
DB	819	AAACACAGAGATGAGAGAAATTCATGAGAGTCCGCAAAATTTGTGTTGAAATATAT	878								
QY	828	AGAAGCGCGAGACAAAACGCTATATATGAGAGAGGAGAGAAATGTAAGAGCCCGCGC	887								
DB	879	TGGAACAAAGAGAGGCTATTAAATATGCGGAAAAACGCAACGACGACATGCTGGGATT	938								
QY	888	GAGATTTTGTGGATTATGATTCAGCAAAAGAAATGACGATTCAGACATTTGGA	947								
DB	939	GTGTTGATGATCAACACGAAAGCAACACAGAACTTAAGATGAGTACAGATATATATGA	998								
QY	948	GAGAGTAAAGCTTTTCTTTCGCGCGGAGAAACAGAACTTCAATCTGAGAGGTGAC	1007								
DB	999	GGAATGCAAGCTATTCTACTTCGAGAGAAATGAGAACAAATCGGTGTCTCATCTGGAC	1058								
QY	1008	GACCATTTTGTATCATGACCCCGAGTGGCAGGCCAAAGCAAGTATAGGTCTCTAG	1067								
DB	1059	GTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1118								
QY	1068	GATGCGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1127								
DB	1119	CCACTTTGAG	1178								
QY	1128	TATGATTTGAAAGAGTCTTAAAGTTGATTCACCAATATGATGATGATGATGATGATG	1187								

Db 1179 AATGATCTATATAGAGGCTTTAGGTTGATACCTCCAGTGGTCTTGTGAAACAGCGAAC 1238
 Qy 1188 TAAATCGGATGTGAGTAGAGGTTACAAATCCATGTGGACGAGCTTCTATCCC 1247
 Db 1239 ATTGAAGGAACGAATCTCGGTGCATCAAAATTTCCAGGTGATATGAACCTATATCTCC 1258
 Qy 1248 AATCATAGCGGTCCATCATGACCAAGCCATTTGGGGTATGACGTGAAGCAATTCATCC 1307
 Db 1299 CATCTCTGTTTATACCAATGATCTGTAATTTGGGCAAGATGCAAGTGAATTCATCC 1358
 Qy 1308 AGCTCGGTTTGGGATGAGTGGCGCGTGTGCAACACCCCGTTGGC---TTCATACC 1364
 Db 1359 AGGTAGAGTTTCTGATGACATCTCCATGACCAAGCAAGTATGATATGCTCTTCTCC 1418
 Qy 1365 GTTTGGCTCGGAGTTGTATCATGATGATGTTGTGCAATCTTGTCTATCTTCAAGCCAAAT 1424
 Db 1419 ATTGGATGGGCTCGTCTGATCTGCAATGGCCAGAGCTTGGCACTTGTGAAGCCAAAT 1478
 Qy 1425 GACACTCGCTGATATGATCCACGCTTCACTTCACTTGGCTCTTATCATGATGC 1484
 Db 1479 GGCTCTGACGATGATCTCCAGCATTTCTATGTAACCTTGGCGTCTCATACCAAGC 1538
 Qy 1485 ACCTACCGTCTTATGTTGCTTATCTTCAACATGTTGCAACCAATCCTTCCGAGATT 1544
 Db 1539 ACCCTATATGATGTTAATCTCCGTCACAGAGGTTGCCAAATTAAGCTGAAGAGAT 1598
 Qy 1545 GACCAATCATGAGA 1559
 Db 1599 TTGACATATCAGA 1613

RESULT 10
 CTRCTP450P 1854 bp mRNA linear PLN 16-JAN-2001
 LOCUS Catharanthus roseus cytochrome P-450 protein (CYP72) mRNA, complete
 DEFINITION cds.
 ACCESSION L10081
 VERSION L10081.1 GI:167483

SOURCE

ORGANISM

Catharanthus roseus (Madagascar periwinkle)
 Catharanthus roseus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Gentianales; Apocynaceae; Rauvolfiaceae;
 Vincaceae; Catharanthus.
 1 (bases 1 to 1854)
 Vetter, H.-P., Mangold, U., Schroeder, G., Warner, F.-J.,
 Weick-Reichert, D. and Schroeder, J.
 Molecular analysis and heterologous expression of an inducible
 cytochrome P-450 protein from periwinkle (Catharanthus roseus L.)
 Plant Physiol. 100, 998-1007 (1992)

REFERENCE

JOURNAL
 TITLE
 AUTHORS
 Immler, S., Schroeder, G., St-Pierre, B., Crouch, N. P., Hotze, M.,
 Schmidt, J., Strack, D., Maren, U. and Schroeder, J.
 Indole alkaloid biosynthesis in Catharanthus roseus: new enzyme
 activities and identification of cytochrome P450 CYP72A1 as
 secologanin synthase
 Plant J. 24 (6), 797-804 (2000)

TITLE

JOURNAL
 MEDLINE
 PUBMED
 20575722
 11135113
 3 (bases 1 to 1854)
 Schroeder, J.
 Direct Submission
 Submitted (27-APR-1993) Biologie II, Universitaet Freiburg,
 Schaezlestr. 1, Freiburg 79104, Germany
 Location/Qualifiers
 1..1854
 /organism="Catharanthus roseus"
 /mol_type="mRNA"
 /db_xref="taxon:4058"
 /note="cell suspension culture"
 1..1854
 /gene="CYP72"

FEATURES

source

gene

CDS

57..1631
 /gene="CYP72"
 /note="putative; CYP72 protein"
 /codon_start=1
 /product="cytochrome P-450 protein"
 /protein_id="AA33106.1"
 /db_xref="GI:167484"
 /translation="MENDMTTIRKAAITPFAVYAMNREVDNMFPRKIERLRQ
 QFRGNPYRLVDVKSQMDHGLSKRMEFNDVLPMPHINHTITGKNSFTM
 MGRIPRIHVEPELIEKVLTHSSKYQNDVNPVLFLLTGVSGFEAGMKERII
 SPATLEKLSMPLPAFALCYHMDLTWMEIAEKQSGHEVDIPITDVLISVAF
 GSVEEGKIFELRKEIMLDTICRVDVLPQMSYLPTRNKRMEINELTDMRFI
 INKMMLKAGEPEDDLGLVLESNIOELQKGNKDGMSINDVIECKLYPAGQ
 ETTGVLTMTTITLLSKRPEWGERAREVIOARGKNRPEREINHLKYVMILYEVL
 YPPIDITFKVHVDITLGSYTIYAGVOWLPYTMHRESISNGEAMENRPFVGV
 ANATKNNVTILPSPMPRVCLGQFALQKLGILNIIORPFPDVAISVHAPFTILT
 VQPFQSHVITYKLES"

ORIGIN

Query Match 18.5%; Score 289.6; Db 8; Length 1854;
 Best local similarity 50.7%; Pred. No. 9.8e-72;
 Matches 793; Conservative 0; Mismatches 744; Indels 27; Gaps 3;

Qy 1 ATGAGGAAGAAGTACAGCTGCTGTTATGTTGAGACCAAGAAGATTGAGAA 60
 Db 57 ATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 116
 Qy 61 AGCTTTATATGTAAGGATATGCTCTGTATGTTGAGACCAAGAAGATTGAGAA 120
 Db 117 ATGGCTTGGGATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 176
 Qy 121 CATTTCTTAAACAGAAATTCAGAGTCTCTTATCATTTCTTATTCGGAATGTAA 180
 Db 177 GCTTAAAGCAGCAAGTTTAAAGGAAATCTTAAAGATCTGTTGAGATGTAAG 236
 Qy 181 GAATTTGTAAGATGATGCTTAAAGCTTCTTCACTCTTAAAGCTTCTTCAATAT 240
 Db 237 GAGATGGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 296
 Qy 241 CTTCCTAAGTCTCTCTTTTACCATCATGAGAGAAATCTAGGCTCTCATTTCCG 300
 Db 297 GTTCTCTCCCTCATGACCAATATTAACACATCATCAATCTTACGTAATTCCT 356
 Qy 301 GTTTGTTCCGTCACATCTTCCGTTAACCGATCCGATCTTATGATGACAGATC 360
 Db 357 ACATGATGGAAGATTCACAAATCATGATGATGATGATGATGATGATGATGAT 416
 Qy 361 TTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 417
 Db 417 TTGACCACTAAGCAAAATCCAAAGAACTTGTATGATGATGATGATGATGATGAT 476
 Qy 418 GAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 477
 Db 477 CTTCACCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 536
 Qy 478 AGCCCTATTTTCAATGAGAGAAATCTTAAATGCTTGTACAGTTGTGGAAGTGG 537
 Db 537 TCCCTGCTTCACTTGAAGAACTTAAATGATGATGATGATGATGATGATGATGAT 596
 Qy 538 ACTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 597
 Db 597 CATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 656
 Qy 598 GTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
 Db 657 ATCTTCTCCACGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 716
 Qy 658 AGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
 Db 717 ACATATGAGAGAGAGCAAAATCTTCACTATGTAAGAAATCTATGATGATGATGAT 776
 Qy 718 GAGCTTTTCAAAAGTCTTCACTCCGCTATGATGATGATGATGATGATGATGATGAT 777

Db 777 GACTGATGAGAGATGCTACATTCAGAGATGAGACTTCTGCAACCAAGAGAGACAG 836
 Qy 778 AACTTCGGAGATGACAAAGAGAGATGAGATGCTGTTGAGAGATGAGAGAGAG 837
 Db 837 AGATGAGAGAGATGACAAAGAGAGATGAGATGCTGTTGAGAGATGAGAGAGAG 836
 Qy 838 AGACAAAGAGATGAGAGAGAGAGATGAGATGAGAGAGAGAGAGAGAGAGAG 837
 Db 897 ATGAGAGAGATGAG 956
 Qy 898 T-----TGGATTAATGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 936
 Db 957 TCAAACTTCAGAGAGATGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1016
 Qy 937 GACATGAGAGAGAGATGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 936
 Db 1017 GAGTATGAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1076
 Qy 997 CTGAG 1056
 Db 1077 CTGAG 1136
 Qy 1057 GAGATGAG 1116
 Db 1137 GAGATGAG 1193
 Qy 1117 AAGATGAG 1176
 Db 1194 AAGATGAG 1253
 Qy 1177 ATGAG 1236
 Db 1254 AAGATGAG 1313
 Qy 1237 CTGAG 1296
 Db 1314 GAGATGAG 1373
 Qy 1297 GAGATGAG 1356
 Db 1374 GAGATGAG 1433
 Qy 1357 TATGAG 1416
 Db 1434 TATGAG 1493
 Qy 1417 GAGATGAG 1476
 Db 1494 GAGATGAG 1553
 Qy 1477 GAGATGAG 1536
 Db 1554 GAGATGAG 1613
 Qy 1537 CGGA 1540
 Db 1614 AAGA 1617

RESULT 11
 CTRP450B 1633 bp mRNA linear PLN 22-APR-1994
 LOCUS CATHARANTHUS ROSEUS Cp3 cytochrome 450 (CYP72C) mRNA, 3' end.
 DEFINITION
 ACCESSION L19075.1 GI:404689
 VERSION cytochrome 450.
 KEYWORDS CATHARANTHUS ROSEUS (MADAGASCAR PERIWINKLE)
 SOURCE CATHARANTHUS ROSEUS
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; Lamiales; Gentianales; Apocynaceae; Rauvolfiaceae;
 Vincaceae; Catharanthus.
 REFERENCE 1 (bases 1 to 1633)

AUTHORS Mangold, U., Eichel, J., Batschauer, A., Lanz, T., Kaiser, T., Spangenberg, G., Merck-Reichardt, D. and Schroeder, J.
 TITLE Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from Catharanthus roseus, and transgenic expression in tobacco and Arabidopsis thaliana
 JOURNAL Plant Sci. 96, 129-136 (1994)
 REFERENCE 2 (bases 1 to 1633)
 AUTHORS Joachim Schroeder
 TITLE Direct Submission
 JOURNAL Submitted (30-SEP-1993) by Joachim Schroeder, Biologie II, Univ. of Freiburg
 COMMENT Schaezlerstr. 1, Freiburg, Germany D-79104
 Original source text: Catharanthus roseus (strain Cp3) cDNA to mRNA.

FEATURES

source

location/qualifiers
 1..1633
 /organism="Catharanthus roseus"
 /mol_type="mRNA"
 /strain="Cp3"
 /db_xref="taxon:4058"
 /cissue_1ib="NM1149"
 1..1633
 /gene="CYP72C"
 <1..1553
 /gene="CYP72C"
 /note="putative"
 /codon_start=3
 /product="cytochrome P450"
 /protein_id="AA17746.1"
 /db_xref="GI:404690"
 /translation="1ATLIFALVAMAMRVLAMVETKRIEKLROGGERNDYREL
 VGVKSGKQHQALSNPWFENDIVPRLMPHINTITVGRNSFTWGR.LPRLHVE
 PELIKELVTHSSYQNFENPPLVLFLLTGVSFEGAKMHRILSPAFLEKLS
 MLPAFPICHSDMLTKMKELAEKSHEDVLPPTVLTSDVLSKVAFSTYDEGKIF
 RLKELMDLIDCMRNVYRFGSVLPTRKRMKEINKEIDMLRFIINKMKALKG
 EPEPDLLVYLESNIOEFOKQGNKDGAMTINVECKLPYRAGGETTGVLMTT
 ILSKHEPQERREVRLOAFGRNKEFERELNHLKYEMILYEVRLRPYIDLTLLI
 HDPTKGPYVITPAGTQVMLPTVLMREKSIWGEDMERNPFRFVGNARXNVTL
 PFGVRCVLGONFALLOKGLAMLIORFEDVAPSVYHAFITLVOPQFSHVLY
 KLERONF"

ORIGIN

Query Match 18.0%; Score 281.2; DB 8; Length 1633;
 Best Local Similarity 51.1%; Pred. No. 2.5e-69;
 Matches 757; Conservative 0; Mismatches 698; Indels 27; Gaps 3;

Qy 83 TGTCTCTGTATGAG 142
 Db 52 TGTCTCTGTATGAG 111
 Qy 143 GAGTCTCTGTATGAG 202
 Db 112 GAGTCTCTGTATGAG 171
 Qy 203 AAGCTTCTGTATGAG 262
 Db 172 AAGCTTCTGTATGAG 231
 Qy 263 AAGCTTCTGTATGAG 322
 Db 232 TTAACCAACACATTAACCAACCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 291
 Qy 323 GGTAAAGGAG 382
 Db 292 GGTAAAGGAG 351
 Qy 383 AG 439
 Db 352 AG 411
 Qy 440 TCAAGGAG 499
 Db 412 TCAAGGAG 471

QY 500 ATCTAGATTGCTTGTACAGTTGTGTGTAAGAGTGTGATATGTTGATTAATGCT 559
 DB 472 AACTTAAGTCAATGTTCCGGCTTTGGCATTATGCTATCAATGATGTTGACGAATGGG 531
 QY 560 CCGATTAAGTATCAGAAAACGGTAAAGTTAGAGTATGCTATGATGGTTTCAATTT 619
 DB 532 AGAATTAGCAGAAAAGAAAGATCCCATGAAGTATATCTTCCCAATTTGATGTTT 591
 QY 620 TGAATGAAGTATTTATTAAGTAAACGCTTTGGAATGATGATGAAGTGTGACAG 679
 DB 592 TGAAGATGATGTAATTTCAAGGTTGCAATTTGTTGACATATGATTAAGAGCAAAA 651
 QY 680 TTTTGCATCTCAAGCTCAACAATGCTTTTGTGCTGAAGCTTTCAAAAAGTCTTCA 739
 DB 652 TCTTCAAGCTTATGAAAAGAACTCATGATCTCATATGATTTGACATGAGATGCTTACA 711
 QY 740 TTCTGGCTATATGATTTTTCGACAAAGAGGAATTTGAAGTCTGGAAGTTAGCAAG 799
 DB 712 TCCAGGGGTGAGGCTACTTGCACAAAGAGAAACAGAGATGAAGAAATTAACAAAG 771
 QY 800 AGATTAAGAAAGTGGTGTGTAAGCTGATAGAGCGGAGACAAAACGCTATAGATGAG 859
 DB 772 AATTAACAGATATGCTAAGTTATCATATCAACAAGGAATGAAGGCTTGAAGGCTGAG 831
 QY 860 AAGGGAAGAATGTAAGAGCCGCGCG-----CGAAGATTGT 898
 DB 832 AACCAAGTGAAGATGATTTGTTGGAGTCTGTTGGAATCAAAATTCAGAAATTCAGA 891
 QY 899 TGGGATTAAGTATCAGCAAAAGATGACGTTGAGACATGTTGAGAGAGTAAAA 958
 DB 892 AACAGGAACAGAAAGATGATGATGATCAATCAATGATGATGATGATGATGATGAT 951
 QY 959 GCTTTTCTTCCGCGGAAAACAGAACTTATATCTGCTGACGTGACGACATCTTGC 1018
 DB 952 TGTCTATCTTGTGCTCAAGAAACCCGAGTTTGTCTCAATGACCAACATTTAT 1011
 QY 1019 TATCATGACCCCGAATGCGACAGCCAAACAGATGATGAGGCTCAGGATCTGCGGT 1078
 DB 1012 TGACCAAGCATCCGAGTGGCAGAGGACGATGAGGAAGATTTCTCAAGCTTTGGC- 1070
 QY 1079 CACGTGATGTCCTTACCAAGAACCATGCTGTTAAGCTTAAACGTTAGATGATCTGA 1138
 DB 1071 --AAGAACAGCCTGAATTTGAACGTTGATCATCTAAATATGTTCTAATGATCTTAT 1128
 QY 1139 ACGAGCTTTAAGTTGATTCACCAATAGTACGATGACGACGCGCTAAATGGGATG 1198
 DB 1129 ACGAGGTTCTAAGTTGATTCACCGCAGTTATGATGCTGACCAAGATTTATTCAGAGACA 1188
 QY 1199 TGAAGTGAAGGCTTACAAATCCCATGCTGACGAGGCTTCTAATCCCATATATACGG 1258
 DB 1189 CAAAGTTAGGCGCTTACAAATCCCTGACAGAAACAAAGTATGTTGCCGACGATATGC 1248
 QY 1259 TCCATCATGACCAAGCCATTTGGGCTAATGACGTAACGATTCATCCAGCTCGGTTG 1318
 DB 1249 TTCAATAGAGAAAGCATTTGGGAGAAAGATGCAATGAAATTCACCCCAATGATTTG 1308
 QY 1319 CGGATGAGGCGCGGCTGCTGCCAAACACCCCGTTGCTCATACGTTTGGCGTGGAG 1378
 DB 1309 CTGATGAGTGGCCAAATGCCCAAGAAACATATGTCATATTTGCCATTCAGTTGGGAGC 1368
 QY 1379 TTCTATCATGATGATGTCAGAACTTGTCTATCTTACGCGCAATTAACCTGCTTAA 1438
 DB 1369 CTAGAGTTGCTCGGCCAAAACCTTTCATTCAGCAAAATTAAGGTTGGCAATGA 1428
 QY 1439 TGAATCAAGCTTCAACCTTCACTGATGCTCTCTATCATGATGACCTACCGTCTTA 1498
 DB 1429 TTTTCAACGCTTCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1488
 QY 1499 TGTGCTTATCTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1540
 DB 1489 TTAAGTTCACCTCAGTTGTTCTCATGATGATGATGATGATGATGATGATGATGATGAT 1530

RESULT 12

AY096748 1570 bp mRNA linear PLN 18-SEP-2002
 LOCUS Arabidopsis thaliana putative cytochrome P450 protein (At3g14690)

DEFINITION

mRNA, complete cds.

ACCESSION

AY096748.1 GI:20465786

VERSION

FLI CDNA.

KEYWORDS

Arabidopsis thaliana (thale cress)

SOURCE

Arabidopsis thaliana

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 1570)

Yamada, K., Ban, H., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,

Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,

Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,

Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,

Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,

Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,

Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A.,

Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Open Reading Frame (ORF) Clones

Unpublished

2 (bases 1 to 1570)

Yamada, K., Ban, H., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,

Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,

Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,

Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,

Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,

Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,

Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A.,

Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Open Reading Frame (ORF) Clones

Unpublished

Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan

Street, Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RFLP cDNAs (RFLP cDNA: 'RIKEN

Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGC (SSP) Consortium members constructed and

sequenced the PUN1 (ORF) clones using the RFLP cDNAs. Yamada, K.,

Ban, H., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,

Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,

Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,

Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,

Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P.,

Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGC) and Seki, M. (RIKEN GSC) contributed equally

to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP

/PGC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis

genome submitted to GenBank.

Location/Qualifiers

1..1570

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="u11126"

/ecotype="Columbia"

/note="This clone is in pUN1 51."

1..1570

/gene="At3g14690"

1..1539

/gene="At3g14690"

/codon_start=1

/evidence=experimental

/product="putative cytochrome P450 protein"

CDS

gene

FEATURES

source

/protein_id="AA020382.1"
 /db_xref="gi:20465787"
 /translation="MEISVASVTISVLAIVSWIWRLLQWVFKPKMLFHYLRQGL
 AGTYPLVVDLKNFTMLSPASKPIKLTDDISPRVVRPLQMFRTYRTYFWPGR
 IPTITMDPEQIKENKVDFOKEPHEPLATIIAGLAVDGDMAKRRILNPAFH
 IEXIKMVPAPHOSCRVYEMDQVSDKSCSEVWPELVMTADVISAFFGSY
 KEGORIPELOLAOLIIQAFKAFIPGYSLPTKSRKMAAREIIVILGIVNKR
 LBARAGBAPSDDLIGLIESNLCOTENAGSTEDLMECKLEYFAGORTSGLVWY
 MYLSQHDWAPARAREVKOVFGDKEDPAQNLQKVMIIIEVRLVLPYTLTRA
 IHKEIKGDLTPGVOISPLILVCHDIELMKNDAEFPPDFKGLSKATKSOVSF
 FPFAPGPRICIGONFALLEAKMMMLILRFSEISVHYATITTHPQGAOLI
 MHKL"
 1540..1570
 /gene="At3g14690"

ORIGIN

Query Match 17.7%; Score 276.4; DB 8; Length 1570;
 Best Local Similarity 51.3%; Pred. No. 6,1e-68;
 Matches 746; Conservative 0; Mismatches 696; Indels 12; Gaps 4;

3'UTR
 90 GTTATGTGGAGACCAAGAAATTGAAGAACTTTCTCTPAACAAGAAATTCGAGTCC 149
 81 GGTGTGTTTAAAGCAAGATGCTTGAACATTAAGTGAAGAAAGCAAGTCTTCGCGAAC 140
 150 TCCCTATCATTTCTTCATCGAAGATGTTAAAGAACTTGTGAATGATGCTTAAAGCTTC 209
 141 ACCTTAACAGCCTCTAGTCGGGATTTGAAGAAAGATTTACCATGTTGTGGAAGGCAAG 200
 210 TTCTATCTTATGCTTCTCTCAACAATTTCTCCAGAGTTCTCTCTTTTACATCA 269
 201 GTCCAAACCCCTCAACTACCGATGACATCTCACTGCTGTCGTCTTATCCCTTGA 260
 270 CTGAGAAATACTACGCGTGTACATTTCTGTTGTTGTTCCGTCAACTTCCGGTTAC 329
 261 AATGTTCAAGACTTACGGAAGACTTACTTAACTGTTGAGCTTACCAACATCAC 320
 330 GGTAGCCGATCTGATTTGATGAGAAATCTTCTTAAGTCTGATTTACGGAAGAA 389
 321 AATTAATGATCTTGACAAATCAAGAAAGTGTTCACAAAGTTATGATTTCCAAAGCC 380
 390 TGAAGTCACTCTTGGTTTAAACAAGTGAAGCGATGACTTACTAGTCTCAAGGTGA 449
 381 GCATACATTTCTTTGGCA---CAATTATAGCTTAAGAGACTCCGTAATATGATGTA 437
 450 AAATGGGCTCATCTGAAATAATCATTAAGCCCTTACTTCAATGAGAAATCTTAAGTT 509
 438 TAAATGGCAAAACACCGAAGATCATCAACCCGCTTTTCAATAGAGAAATCAAGAA 497
 510 GCTTGACCAAGTTGTTGAAGTGTGATGATGTTGATGATTAATGATGTCGATTAAGTT 569
 498 TATGTAAGCTGCTTCCACCAAGCTGTAAGGAATTTGTGCGAATGGAATCAAGTATG 557
 558 ATCGATTAAGAGGTGCTCATGAGGTAGACGTTGGCTGGCTTGTGATATAGACTGC 617
 570 ATCAGAAAACGGTGA---AGTTAGAGTATGCTATAGAGTGTTCAGATTTTACTGA 626
 627 AGATGTTATTTAGTGAACAGCTTTTGAAGTGAATGATGAATGATGAGAGTATTTGCG 686
 618 AGATGATCTCTCGTACGCAATTTGGCGAGCTTCAAAAGAGGCAAGAGATATTTTGA 677
 687 ACTTCAAGCTCAACAATGCTTCTTGTGCTGAAGCTTTCAAAAGTCTTCAATCTCG 746
 678 GCTCCAGAGAACTGCAAGAGTATCATTTAAGCTTTTGAAGAGCTTTCATCCCTGG 737
 747 CTATAGATTTTTCGACAAAGAGAAATTTGAAGTCTCGGAATTTAGCAAGAGATAG 806
 738 ATATAGTATCTCCCAACAAAATGATACAGAAAGAAATGAAGACGACGCAAGAAATCA 797
 807 GAAGTCGTTGTAAGCTGATAGAGCGCGAGAACAAAACGCTATATGATGAGAAAGGGA 866
 798 AGTTATCTAAGAGGATCTTAACAAAAGTTTAAAGGCGCAAGAGCTGATGAAGCACCC 857
 867 AGATGTAAAGACCCGCGCAAGATTTGTTGGATTAATGATTCAGCAAGAAAGAT 926

858 AAGCGAATTTGTGTGGAAATCTTCTGTGATGAATTTGAGCAAAACGAAAGAAACGG 917
 927 GAGGCTT---CAGACATTTGAGAGAGATTAAGAACTTTTCTTCCCGGGAACAGAC 983
 918 AATAGTACGAGATCTGATGAGAGATGACAGTCAAGTTGTTCTATTTTCCCGGGAAGAGAC 977
 984 AACTTAAATCTGCTGACGTGACGACCAATCTTGTATCCATGACACCCGAGTGGACAGGC 1043
 978 AACATCAGTACTTCTGCTTGGACAAATGATGCTTTAAGCAACACCAAGATTTGGACG 1037
 1044 CAAGCCAGTATGATGCTTCAAGGCTGCGGCTGACGATGCTCCCTACCAAGACCA 1103
 1038 TGTGACGAGAAAGATGAAGCAAGTTTGGCA---TAAAGAACTGATGCAAGAG 1094
 1104 TGTGTTAAGCTTAAACGTTGATGATGATCTTGAACAGTCTTTAAGTTATCCACC 1163
 1095 ACTCAACAGCTCAAGATTAATGATGATGATATATATAGAGTCTTAAAGCTATATCTTC 1154
 1164 AATGTAGCTACATTTGACGCGCTAATGCGATGATGATGATGATGATGATGATGATGAT 1223
 1155 AGTAAACCAAGCTGACGAGCATTTCAAAAGAGCTGAAAGCTGAAAGCTGATGACCTTAC 1214
 1224 ATGTGACAGGAGCTTCAATCCCAATCATAGCGCTCATATGACCAAGCAATTTGGGG 1283
 1215 AGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274
 1284 TATGACGTGAACGAATTAATCCAGCTCGGTTGCGAGATGAGAGTCCGCTGCTGCA 1343
 1275 AAACGACGACAGCAAGTCAATCTTGAACAATTAAGATGATGATGATGATGATGATGAT 1334
 1344 AACCCGCTGCTTCAATACGTTGCTGCGAGTCTGATGATGATGATGATGATGATGATGAT 1403
 1335 GAGCAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1394
 1404 TGTATACTTCAAGCCAAATGACATGCTGCTGATGATGATGATGATGATGATGATGATGAT 1463
 1395 TGCCCTGTTAAGGCAAAATGCAATGAGCAATGATGATGATGATGATGATGATGATGATGAT 1454
 1464 GCTTCTTCTTCAAGATGACATGACCTGATGATGATGATGATGATGATGATGATGATGAT 1523
 1455 TTCTCTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1514
 1524 ACCAATCACTTCC 1537
 1515 TGAAGTTATGATG 1528

RESULT 13
 AY050827
 LOCUS
 DEFINITION
 mRNA, complete cds.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Arabisopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Fossids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 1692)
 Yamada, K., Liu, S.-X., Sakano, H., Pham, P.-K., Ban, J., Chung, M.-K.,
 Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
 Lin, J., Miyake, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
 Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K.,
 Davis, R.W., Ecker, J.R. and Theologis, A.
 Arabidopsis Full Length cDNA clones
 Unpublished
 2 (bases 1 to 1692)
 Yamada, K., Liu, S.-X., Pham, P.-K., Ban, J., Banjo, F., Dale, J.M.,

TITLE
JOURNAL
COMMENT

Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carrinci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamuya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koeseema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Tracy, S.E., Shinohara, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission
Submitted (08-AUG-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamuya, A., Sakurai, T., Carrinci, P., Kawai, J., Hayashizaki, Y. and Shinohara, K.

The Salik, Stanford, PGRS (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada, K., Liu, S.X., Pham, P.K., Ban, J., Banno, F., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koeseema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGRS) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinohara, K. (RIKEN GSC) and Theologis, A. (SSP/PGRS) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES
source

Location/Qualifiers
1..1692
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAPL08-11-K16 (R11126)"
/ecotype="Columbia"
/note="this clone is in a modified pbluescript vector (FLC-1) as a BamHI/XhoI insert."

gene
1..1692
/gene="At3g14690"
1..31
/gene="At3g14690"
32..1570
/gene="At3g14690"

CDS
/codon_start=1
/evidence=experimental
/product="putative cytochrome P450 protein"
/protein_id="AAK2762.1"

/db_xref="gi:15292787"
/translation="MEISVASVTISVLAIVSWIWRITLQWVFKPKLLEHLEKQGL
AGTYPLVGLKKNFTLSKSKPLKLDSPVIVYPLQWMTKRYFTFTFGP
IPTTIMPEOIKKENVVDPOKPHFPLATIIAGLAVGDMKAGRIINPAFI
IEIKIMVPAHOSCEVGVMDQVSSCEVAVGDMVSMADVISIATAGSSY
KESORIELOEACOLIOAFRAPIRGYSALPTKSNREMKAAAEIOVIRGVNVR
IRABAGABEDDLGITLESNIRPTEGNSMREDMLECKIFPACGETTSVLTWT
VWLSQHDWGRARAREKOVFGKSDGAGNQLKVTMILYEVLRIPVTLQTRA
IHKLKGLDITLPGVQISLPIIVQHDLEWGNDAEINPDRFGDSKATKQVSP
FEPFAGPRICIGNPFALLEKAMALILRFSPESIVYAPVITIHPOFAGOLI
MKKL"

3' UTR
1571..1692
/gene="At3g14690"
1676
/gene="At3g14690"
/note="not present in genomic sequence"

ORIGIN
Query Match 17.7%; Score 276.4; DB 8; Length 1692;
Best Local Similarity 51.3%; Pred. No. 6.2e-68;
Matches 746; Conservative 0; Mismatches 696; Indels 12; Gaps 4;

Qy 90 GTTATGTCGAGACCAAGAAAGATTGAAGAATCTTCTTAACAGAAATTGAGATCC 149
Db 112 GATTGGTTTAAAGCAAGATGCTGAGCATACCTAGAAAGCAAGCTTCCGGAAC 171
Qy 150 TCCCTATCATTTCTTCATCGAAATGTTAAAGAACTTGTGAATGATGCTTAAAGCTTC 209
Db 172 ACCCTTACAGCCCTTAGTCGAGATTGAAGAAATTTCAACATGTTGCGAGGCAAG 231
Qy 210 TTTCTATCTATAGCTTCTCTCAATATTTCTTCCNAGTCTCTCTTTTACATCA 269
Db 232 GTCCAAACCCCTCAAACTTACGAGATGACATCTCCTGATGCTGCTATCTCTTCA 291
Qy 270 CTGAGAAATCTACCGTGTACATTTCTGTTTGTGCTCAACTTCCGTTAAC 329
Db 292 AATGTTCAAGACTTACGAGAGACTTACTTACATGTTTGAACCTATCCACATCAG 351
Qy 330 GGTAGCCATCTGATTTGATCAGAGATCTTCTTAATCTGAATTCAGAGAGAA 389
Db 352 AATATGATCTCTGACAAATCAAGAAAGTTTACAAAGTTTATGATTTCAAAAGCC 411
Qy 390 TGAAGCTACCTTGTGTTAAACAATTGAAGCGATGACATTAATGCTCAAGTGA 449
Db 412 GCATCATTTCTTTGGCCA---CAATTAATGTTAAAGACTGCTAATTAATGATGTA 468
Qy 450 AAAATGGCTATCATGCAAAATCATTAAGCCCTACTTTTCAATGAGAAATTTAAT 509
Db 469 TAAATGGGAAAACCGAAGAAATCATACCCGCTTTTCAATGAGAAATTTAATG 528
Qy 510 GCTTGTACCAATGTTGTTGAAGATGATGATGATGATGATGATGATGATGATGAT 569
Db 529 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588
Qy 570 ATCAGAAACCGTGA---AGTTAGGTAGATGATGATGATGATGATGATGATGATGAT 626
Db 589 ATCGATTAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 648
Qy 627 AGATGTTATGATGAACAGCTTTTGAAGTATGATGATGATGATGATGATGATGATGAT 686
Db 649 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708
Qy 687 ACTTGAAGCTCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 746
Db 709 GCTTCAAGAGAACTAGCAAGCTCATCATCATCATCATCATCATCATCATCATCAT 768
Qy 747 CATAGATTTTTCGCAAGAGGAAATTTGAAGTCTCGAAGTTAGACAGAGATTAAG 806
Db 769 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 828
Qy 807 GAAAGCTGTTTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 866
Db 829 AGTTATCTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888
Qy 867 AGAATGTAAGAGCCCGGCGGAAAGATTTGTTGGATTATGATTCAGGCAAGATGT 926
Db 889 AAGCAGCATTTTGTGGGAATTAATCTTGAATGCAATTTAGGCAAGAGAGAAACG 948
Qy 927 GAGCGT---CAGCAATTTGAGAGAGTGAAGTAAAGTTTCTTCCGCGGAAACAG 983
Db 949 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008
Qy 984 AACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1043
Db 1009 AACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
Qy 1044 CAAGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1103
Db 1069 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
Qy 1104 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1163
Db 1126 ACTCAACAGCTCAAAAGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1185

QY 1164 AATGTAAGCTACGATTGAGCGGCTAATTCGATGTGAAGCTTGAAGGGTACAAATCC 1223
 DB 1186 AGTAAACCAAGCTACCTGAGCCATTCACAAAGAGCTGAAGCTGAGCTTGAACCTTAC 1245
 QY 1224 ATGTGACGAGAGCTTCAATCCATTCATAGCGCTCATCATGACCAAGCCATTTGGGG 1283
 DB 1246 AGCGGTGTTCAGATCAGTCTACCTATTCGTGAGTCCAAACGACATCGAGCTGTGGG 1305
 QY 1284 TAAAGAGTGAAGCAATTCATCCAGCTGGTTGGAGTGAAGTGGCCGCTGTCCAA 1343
 DB 1306 AAGACGACGACAGAGTCAATCTGACAGATTTAAAGAGTCTTTCAAAAGGCAAAA 1365
 QY 1344 ACACCCCGTTGGCTTCATACCGTTGGCCCTCGAGTCTGTACATGATGTGCAAAATCT 1403
 DB 1366 GAGCAAGTCTCTCTTCCCTTGGGGGAGCGAGATATGCAATAGCGCAAACTT 1425
 QY 1404 TGCTAATCTGAGCCAAATTCAGACTCGTGAATGATCCAGCTTCACCTTCACTT 1463
 DB 1426 TGCCCTGTAGAGGCAAAATGCAATGCGATGATTCGCGAGGTTCTGTTTGAAT 1485
 QY 1464 GGTCTACTATTCAGAGTCACTACCGTCTTATGTTGCTTATCTCAACATGTGTC 1523
 DB 1486 TTCTCTTCCATTCATTCAGCAGCACAACAGTCAATCACTTCAACCAAGTTCGGTGC 1545
 QY 1524 ACCTATCACTTCC 1537
 DB 1546 TCAGCTTATCATGC 1559

RESULT 14
 AB047400 1572 bp mRNA linear PLN 21-MAR-2002
 LOCUS Oryza sativa (japonica cultivar-group) Rice2 mRNA for cytochrome
 DEFINITION P450, complete cds.
 ACCESSION AB047400
 VERSION AB047400.1 GI:18700294
 KEYWORDS

SOURCE

ORGANISM Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Scrophophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriactroideae; Oryzae; Oryza.

REFERENCE

1 Imaishi, H., Shimizu, M. and Ohkawa, H.
 Cytochrome P450 in rice
 Published Only in Database (2002)
 JOURNAL 2 (bases 1 to 1572)
 REFERENCE Imaishi, H., Shimizu, M. and Ohkawa, H.
 Direct Submission
 Submitted (18-AUG-2000) Hiromasa Imaishi, Kobe University, Faculty
 of Agriculture, Rokkodaicho 1-1, Kobe, Hyogo 657-0013, Japan
 (E-mail: hmaishi@kobe-u.ac.jp, Tel: 81-78-803-5945,
 Fax: 81-78-803-5947)

FEATURES

source

1..1572 Location/Qualifiers
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiivar="Nipponbare"
 /db_xref="taxon:39347"
 1..1572
 /gene="Rice2"
 1..1572
 /gene="Rice2"
 /codon_start=1
 /product="cytochrome P450"
 /protein_id="BA85117.1"
 /db_xref="GI:18700294"
 /translating="MLM,GAASQWILAAAAAVALIWLAVSLTWMWTPRRER
 ALRAQIRGNRYLFTGDYENRYLREARKKDPLGCHDIIIPVLPMFSKAEHKK
 PSFTWGPETPRWIMSPESIRVSNKPFHYKSPPTLGLKLGAVSYSGEKNAH
 RLINLPAFHEKXIKMLPVSNCTEMTMRMNSMISIGSEVWVMPFONLGDVLS
 KTAFGSSYREGRIIFOLAESAPRIICARTTFTGVPFLTKXNRLLREIEVSLT
 DRGIIIGRERAIKNGETSNBDLIGLIVSNRSGKALEMTTDEILIECKLPTPAG

ORIGIN

Query Match 17.5%; Score 273.2; DB 8; Length 1572;
 Best Local Similarity 51.2%; Pred No 5,1e-67;
 Matches 750; Conservative 0; Mismatches 698; Indels 18; Gaps 4;

94 TGTGTGAGACCAAGAAAGATTGAAGAACATTTCTTAAACAAAGAAATTCAGTCTCTCT 153
 DB 106 TGTGTGAGACCGCGCGCGCTGAGCGCGCCCTCCGCGCGCAGAGGCAATCCGCGCAACGG 165
 QY 154 TATCATTTCTTCATCGGAAGATTTAAAGAACTTGTGAATGATGCTTAAAGCTTCTCT 213
 DB 166 TACGCTCTTCAACCGGACAGTCCGGAAGACGTCGGCTCAACCGGAGCCGGAAG 225
 QY 214 CATCTATGCTTT--CTCTCAATATTCCTCTAGAGTCTCTCTTTTACATCAC 270
 DB 226 AAGCGCTGCGCGCTCGGCTCCACGACATATCCCGCGGTGCTGCCATGTTTCAGCAA 285
 QY 271 TGGAGAAATCTAGCGGTGACATTTCTGTTGTTGGTGGTCAACTTCCGCTTAAG 330
 DB 286 GCCGTGAGAGACCGGAAACCATCTTACTGTTGTTGGCCCAACGCGCAAGATGATG 345
 QY 331 GTAGCGATCTGTATTTGATCAAGAGATCTTCTAAGTCTGAGTCTTACAGAAAGAT 390
 DB 346 ATTCAGACCTGTAGATCAATAGGAAATGATATCTTAATGATTTGGCCACATATGCAAA 405
 QY 391 GAAGCTACCTTTGTTAAACAACTGAAGGAGATGACTACTTATGCTCAAGAGTGA 450
 DB 406 CCAAGCTTACCGCTTCGGAAGATGCTGACCTCCGAGTTGTAAGCTGAAGCGAG 465
 QY 451 AATAGGCTCATCATCGAAATCATTTAGCCCTACTTTTCATATGAGAAATCTTAAAGTT 510
 DB 466 AATAGGCTCAAGACCGGAGAAATCTGAAATCTGCTTTCACCAACAGAAATTAAGCG 525
 QY 511 CTGTACAGATGTTGTTGAAGAGTGAATGATGATGATGATGATGATGATGATGATGAT 570
 DB 526 ATGCTGCAATTTTCTTAACTGCTGACGGAATGTTTCAAGATGGAAGATTTCAATG 585
 QY 571 TCAGAAACGCTGAAGTTGAGTATGCTATGAGTGTTCAGATTTTGACTGAAGAT 630
 DB 586 TCTATTTGAAGATGTCAGAGATGATGTTGCTGAGATTCAGATTTTCAAGAGAT 645
 QY 631 GTTATTTAGACAGCTTTTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 690
 DB 646 GTCAATTTCAAGACAGCTTCGTAAGATGATGATGATGATGATGATGATGATGATGAT 705
 QY 691 CAAGCTCAAAATGCTCTTCTTGTCTGAAGCTTTTCAAAAGCTTCACTTCTGCTAT 750
 DB 706 CAAGCAGAGTCAAGCAGATATGCAAGCTTTTCAAGCAATTTTTCACAGAGAT 765
 QY 751 AGATTTTTCGACAGAGGGAATTTGAATCTCGAAGTTTGAACAGAGATTAAGAA- 809
 DB 766 TGTGTTTCAACCACTTAAACACAGAGGTTTGAAGAAATTTGAAGAGAGTTCAGCAA 825
 QY 810 -----GTGTTGTTGAAGCTGATGAGAGCGGAGACAAACGCTTATGATGAGA 861
 DB 826 CTTTACAGAGAAATTTGAAAGAGAGCGGCTTATTAATAATGTAACCAAGTAAT 885
 QY 862 GAGAGAGATTAAGAGAGCGGCGCAAGATTTTGTGAGATTAATGATTAAGCAAG 921
 DB 886 GGTGCTTGTGGCTTATTTGTGATGATCAAAATATGAGGAGCAATGGAAGAGAA 945
 QY 922 AATG--TGACGTTTCAAGACATTTGAGAGATTAAGATTTTCTTCCCGGAAA 978
 DB 946 CTAGAGATTAAGAGAGCAATTTTGAAGAAATGCAAGCTATTTTATTTTGAAGAAATG 1005
 QY 979 CAGACAACTTTATCTGCTGACGTGAGACACATCTTGTATCAATGACCGCGAGTGG 1038
 DB 1006 GAGACAACTATGATTTGCTTCTTGAACATTAATTTGTGATGATGACCGCAGAGTGG 1065

Mon May 24 08:10:07 2004

us-09-992-901-1.rge

Page 21

```
QY 154 TATCATTTCTTCATCGGAATGTAAAGAACTTGTGAATGATGCTTAAAGCTTCTCT 213
Db 307 TACCGCTCTTCCACCGCGACGTCCCGAGAAAGTCGGCTCAACCGGAGGCCCGAG 366
QY 214 CATCTATGCTTT--CTCTCAATATTTCTTCTAGAGTTCTCTTTTTCACATAC 270
Db 367 AAGCGCTGCGCTGCGCTCCAGACATATCCCGCGCTGCTCCATGTTACGAAA 426
QY 271 TGGAGAAAATCTACGCTGCTACATTTCTGTTGGTTGGTCCAACTTTCGGTTAAG 330
Db 427 GCCCTGAGGAGCGCGGAGAACATATTCATCTGTTGGTCCCAACCGCAAGAGTATG 486
QY 331 GTAGCCGATCTCTGATTTGATCAGAGATCTCTTCTAGTCTGAGTTCTTACAGAGAA 390
Db 487 ATTTCAGACCTGATCAATATAGGAGATATGCTATATATGTTGGCCACTATGGCAA 546
QY 391 GAAGCTCACCTTGTGTTAAACAATTGAAGGAGATGACCTACTAGTCTCAAGGTGAA 450
Db 547 CCAAGGCTACCCGCTCGGAAAGTGTAGCTTCGAGGTGATGAGTATGAGGCGAG 606
QY 451 AATGGGCTCATCTGAAAAATCATTAAGCTTCTTCAATGAGAAATCTTAAGTTG 510
Db 607 AATGGGCAAGACCGGAGAAATTCGAAATCTGCTTTCACACGAGAAATTAAGCG 666
QY 511 CTGTGACCACTTGTGTTGAAGAGTGTGATGATGATGATGATTAATGTCGATTAAGTTA 570
Db 667 ATGCTGCAATTTTCTTCTAATCTGCTCAAGGAAATGTTTACAGATGAGAAATTCATG 726
QY 571 TCAGAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
Db 727 TCTATGAGAAATGTCAGAGGATGATGATGATGATGATGATGATGATGATGATGAT 786
QY 631 GTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 690
Db 787 GTCATATCAAGACAGATTCGCTAGACCTTATGAGAGAGAGAGAGAGAGAGAGAGAG 846
QY 691 CAAGCTCAACAAATGCTTCTTGTGCTGAGCTTTCAAAAAGTCTTCAATCTGCTAT 750
Db 847 CAAGCAGATCAGCCGACGATTAATACAGCTTTCGACATTTTATACAGATAT 906
QY 751 AGATTTTTCGACAGAGAGAGATTTGAAGTCTCGAAGTATGACAGAGAGATTAAGAA- 809
Db 907 TGTCTTCACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 966
QY 810 -----GTCTGTTGAAGCTGATGAGCGCGAGACAAAGCTATGATGAGAA 861
Db 967 CTCTACAGAGAAATTAATGAAAGAGAGCGGCTATTAATAATGATGAACAGATAT 1026
QY 862 GGGGAGAGATGTAAGAGCGCGCGAGAGATTTGTTGGATTAATGATTAAGGCAAG 921
Db 1027 GGTGACTTGTGGCTTATTTGTGAGTCAATATAGAGAGTCAATAGGAGAGAGAGAG 1086
QY 922 AATG--TGACGTTCAAGACATTTGAGAGAGTAAAGCTTTTCTGCGCGGAAA 978
Db 1087 CTAGGATGATCTAGAGAGAAATTAATGAGGATCAAGCTATTTATTTTTCAGGATG 1146
QY 979 CAGCAACTTCTAATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1038
Db 1147 GAGGCAACATCAATATGCTCACTTGAATTAATTTGTGATGATGAGAGAGAGAGAG 1206
QY 1039 CAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1098
Db 1207 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1263
QY 1099 GACCATGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 1158
Db 1264 GATAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 1323
QY 1159 CACCAATTAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 1218
Db 1324 CCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1383
```

Search completed: May 22, 2004, 19:00:48
Job time : 6292 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2004, 10:50:51 ; Search time 681 Seconds
(without alignments)
9750.270 Million cell updates/sec

Title: US-09-992-901-1

Perfect score: 1563
Sequence: 1 atgaggaagaagatgacag.....tgaccatcatgagatcga 1563

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: geneeqn1980s:*\n2: geneeqn1980s:*\n3: geneeqn2000s:*\n4: geneeqn2001as:*\n5: geneeqn2001bs:*\n6: geneeqn2002s:*\n7: geneeqn2003as:*\n8: geneeqn2003bs:*\n9: geneeqn2003cs:*\n10: geneeqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1563	100.0	3 AAA99072	AA99072 Arabidops
2	318.8	20.4	6 ABR50985	AB50985 DNA encod
3	270.4	17.3	6 ABL60758	AB60758 Malze cyt
4	269.2	17.2	6 ABR50986	AB50986 DNA encod
5	265.4	17.0	6 ABR50983	AB50983 Zea mays
6	241.2	15.4	6 AB212287	AB212287 Arabidops
7	241.2	15.4	6 AAC43009	AAC43009 Arabidops
8	206.2	13.2	6 ADA48311	ADA48311 Rice gene
9	194.8	12.5	6 AB212574	AB212574 Arabidops
10	185.4	11.9	6 AAC41990	AAC41990 Arabidops
11	184	11.8	3 AAA99071	AA99071 Arabidops
12	153.6	9.1	6 ABRN8320	ABN8320 Arabidops
13	141.8	9.1	8 ADA48675	ADA48675 Rice gene
14	104.2	6.7	3 AAC54724	AAC54724 Arabidops
15	102.6	6.6	3 AAC53115	AAC53115 Arabidops
16	100.6	6.4	6 ABR50982	AB50982 Zea mays
17	93	6.0	6 ABR61908	AB61908 Arabidops
18	88.6	5.7	6 ABL93336	AB93336 Arabidops
19	88.6	5.7	6 ABR60704	AB60704 Arabidops
20	77.4	5.0	7 ABR61927	AB61927 Arabidops
21	75.8	4.8	3 AAA68008	AA68008 Eucalyptu
22	75.8	4.8	3 ADD41758	ADA41758 Commarate
23	73.4	4.7	3 AAA68076	AA68076 Eucalyptu

24	73.4	4.7	398	9 ADD41826
25	72.2	4.6	422	7 ABX62213
26	72.2	4.6	424	7 ABX61377
27	70.6	4.5	1392	4 AAS02937
28	70.6	4.5	3150	4 AAF54832
29	70.6	4.5	3150	4 AAF62668
30	70.6	4.5	3150	5 AAF63324
31	70.6	4.5	3225	7 ABZ70676
32	70.6	4.5	4957	7 AAL51096
33	70.6	4.5	4957	7 AAL54548
34	69	4.4	3150	3 AAA46648
35	69	4.4	3150	4 AAD02365
36	69	4.4	3150	7 ACC48000
37	68.6	4.4	285	7 ABX84315
38	63	4.0	1936	9 ADB59019
39	63	4.0	1936	9 ADB53777
40	60.6	3.9	180	6 ABRN75725
41	57.2	3.7	1494	6 ABZ13577
42	57.2	3.7	1494	6 ADA68047
43	55.2	3.5	497	8 ADA49004
44	53.8	3.4	1976	6 ABR48198
45	53.4	3.4	631	6 ABQ66106

ALIGNMENTS

RESULT 1
ID AAA99072 standard; DNA; 1563 BP.
XX AAA99072;
XX 18-JAN-2001 (first entry)
XX Arabidopsis thaliana basl nucleotide sequence SEQ ID NO:1.
XX Arabidopsis thaliana; basl; promoter; cytochrome P450; CYP72B1; plant;
XX brassinosteroid signalling; brassinosteroid synthetase; brassinolide; ds.
XX Arabidopsis thaliana.
XX OS
XX WO200055302-A2.
XX 21-SEP-2000.
XX 16-MAR-2000; 2000MO-US006915.
XX 16-MAR-1999; 99US-0124570P.
XX 14-DEC-1999; 99US-0170931P.
XX 20-DEC-1999; 99US-0172832P.
XX (SALK) SALK INST BIOLOGICAL STUDIES.
XX Neff MM, Chory J;
XX WPI; 2000-638195/61.
XX P-PDSB; AAB23917.
XX Transgenic plants having modulated brassinolide synthesis resulting in
XX insect resistance, dwarfism and darker-green foliage compared with wild-
XX type plants, have nucleic acid encoding Basl polypeptide in its genome.
XX
XX Disclosure; Fig 1A; 104pp; English.
XX The present invention describes a genetically modified plant (1)
XX comprising at least one exogenous nucleic acid sequence encoding a Basl
XX polypeptide, homologous or functional fragment, in its genome or at least
XX one regulatory sequence that modulated expression of endogenous basl gene,
XX homologous or functional fragment, and which is characterised as having
XX modulated brassinolide activity or synthesis. The basl gene encodes a
XX cytochrome P450 (CYP72B1), which has a role in brassinosteroid signalling
XX or synthesis. Overexpression of the basl gene in plants causes a dark

CC green, dwarf phenotype which mimics plants that have low levels of the
CC plant hormone, brassinolide. Overexpression of the *bax1* gene also
CC increases resistance to insects in plants. The present sequence
CC represents the *Arabidopsis thaliana* nucleotide sequence, which is used in the
CC exemplification of the present invention

Sequence 1563 BP; 424 A; 310 C; 377 G; 452 T; 0 U; 0 Other;

Query Match	100.0%	Score 1563;	DB 3;	Length 1563;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1563;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGAGAGAAAGAAAGTACAGCTGCTGATCATTCGAAAGTCTGTGCTGTAACTTAA	60
Db	1	ATGAGAGAAAGAAAGTACAGCTGCTGATCATTCGAAAGTCTGTGCTGTAACTTAA	60
QY	61	AGTCTGTAAATAGTGAAGGGTATGTCTCTGTTATGTGTGAGACCAAGAAAGATTGAAGA	120
Db	61	AGTCTGTAAATAGTGAAGGGTATGTCTCTGTTATGTGTGAGACCAAGAAAGATTGAAGA	120
QY	121	CATTCTCTAAACAAAGGAATTGAGAGTCCCTTATCATTTCTTCAATCGGAATGTATAA	180
Db	121	CATTCTCTAAACAAAGGAATTGAGAGTCCCTTATCATTTCTTCAATCGGAATGTATAA	180
QY	181	GAACTTGTGAGATGATGCTTAAAGCTTTCATCTCATCCATAGCCCTTCTCCAAATATT	240
Db	181	GAACTTGTGAGATGATGCTTAAAGCTTTCATCTCATCCATAGCCCTTCTCCAAATATT	240
QY	241	CTTCCTAAGAGTCTCTCTTTTACATACATCGAGAGAAAATCTACAGGTGTAATATTTCG	300
Db	241	CTTCCTAAGAGTCTCTCTTTTACATACATCGAGAGAAAATCTACAGGTGTAATATTTCG	300
QY	301	GTTTGGTTGCGATCCAACTTTCGAGTTAAACGTAAGCCATCTCTGATTTGATCAGAGATC	360
Db	301	GTTTGGTTGCGATCCAACTTTCGAGTTAAACGTAAGCCATCTCTGATTTGATCAGAGATC	360
QY	361	TTCTCTAAGTCTGAGTTCTTACAGAGAAAGATGAGTCACTCCCTTGGTTAAACAATTGAA	420
Db	361	TTCTCTAAGTCTGAGTTCTTACAGAGAAAGATGAGTCACTCCCTTGGTTAAACAATTGAA	420
QY	421	GGGATGGAATACTTATGCTCAAAAGGTGAAAAATGGGCTCATCATCAAAAATCATTTAC	480
Db	421	GGGATGGAATACTTATGCTCAAAAGGTGAAAAATGGGCTCATCATCAAAAATCATTTAC	480
QY	481	CCTACTTTCAATGAGAAATCTTAAAGTTGCTTTGACAGTTGCTTGAAGAAGTGAAGT	540
Db	481	CCTACTTTCAATGAGAAATCTTAAAGTTGCTTTGACAGTTGCTTGAAGAAGTGAAGT	540
QY	541	GATAGGTGATAAAGTCCGATAGATATCAGAAAAACGTTGAAGTGAAGTGAATGTC	600
Db	541	GATAGGTGATAAAGTCCGATAGATATCAGAAAAACGTTGAAGTGAAGTGAATGTC	600
QY	601	TATGAGTGTTCAGATTTTGACTGAGATGTTATTATGAGAACAGCTTTTGAAGTAC	660
Db	601	TATGAGTGTTCAGATTTTGACTGAGATGTTATTATGAGAACAGCTTTTGAAGTAC	660
QY	661	TATGAGATAGTGTGAGACGATTTTGCAGCTTCAAGTCAACAAATAGCTTCTTGAGCTGAA	720
Db	661	TATGAGATAGTGTGAGACGATTTTGCAGCTTCAAGTCAACAAATAGCTTCTTGAGCTGAA	720
QY	721	GCTTTCAAAAAGCTTCAATCTCGCTATPAGATTTTTCGACAAAGAGGAATTTGAAG	780
Db	721	GCTTTCAAAAAGCTTCAATCTCGCTATPAGATTTTTCGACAAAGAGGAATTTGAAG	780
QY	781	TCTCGGAAGTTTACCAAGAGATTAAGAAAGTGGTTGTAAGCTGATAGAACGGCGAGAA	840
Db	781	TCTCGGAAGTTTACCAAGAGATTAAGAAAGTGGTTGTAAGCTGATAGAACGGCGAGAA	840
QY	841	CAAAACCTTATGATGAGAGAGGGAAGATGTAAAGAGCCGGCGCCAGATTTGTG	900
Db	841	CAAAACCTTATGATGAGAGAGGGAAGATGTAAAGAGCCGGCGCCAGATTTGTG	900
QY	901	GGATTTAATGATCAGGCAAGATGTGACCGTTCCAGACATTTGTGAGAGATGTAAAGC	960

Db	901	GGATTATGATTCAGGGCAAGAAATGACGGTTCAGAGACATTGTGAGAGAGTAAAGC	960
Qy	961	TTTTTCTTCGCGGGGAAAACAGACAACCTTCTAATCTGCTGACGTGACGACCATCTTGCTA	1020
Db	961	TTTTTCTTCGCGGGGAAAACAGACAACCTTCTAATCTGCTGACGTGACGACCATCTTGCTA	1020
Qy	1021	TCCATGACACCCGAGATGCGACGGCCAAAGCACGATGAGAGTCTCAAGGTCGCGGCTCA	1080
Db	1021	TCCATGACACCCGAGATGCGACGGCCAAAGCACGATGAGAGTCTCTCAAGGTCGCGGCTCA	1080
Qy	1081	CGTGATGTCCTTACCAAGACACCATGTGCTTAAGCTTAAACGTTGAGTATGATCTTGAAC	1140
Db	1081	CGTGATGTCCTTACCAAGACACCATGTGCTTAAGCTTAAACGTTGAGTATGATCTTGAAC	1140
Qy	1141	GAGTCTTTAAGTTGTATCCACCAATATGATAGTACGATTGACGGGCTAAATCGGATG	1200
Db	1141	GAGTCTTTAAGTTGTATCCACCAATATGATAGTACGATTGACGGGCTAAATCGGATG	1200
Qy	1201	AAGCTAGAGAGGTACAAATCCATGTCGACGAGCTTCTAATCCCAATCATAGCGGTC	1260
Db	1201	AAGCTAGAGAGGTACAAATCCATGTCGACGAGCTTCTAATCCCAATCATAGCGGTC	1260
Qy	1261	CATCATGACCAAGCCATTGGGGGTATGACGTGAACGATTCATATCCAGCTGGTTGCG	1320
Db	1261	CATCATGACCAAGCCATTGGGGGTATGACGTGAACGATTCATATCCAGCTGGTTGCG	1320
Qy	1321	GATGAGATGCGCGGTGCTGACCAACACCCCGTTGAGTATCCGTTTGAGCTCGAGGT	1380
Db	1321	GATGAGATGCGCGGTGCTGACCAACACCCCGTTGAGTATCCGTTTGAGCTCGAGGT	1380
Qy	1381	CGTACATGATGTGTGAGATCTTGTCTATATCTTGAGGCCAAATTGACATTCGCTGTATG	1440
Db	1381	CGTACATGATGTGTGAGATCTTGTCTATATCTTGAGGCCAAATTGACATTCGCTGTATG	1440
Qy	1441	ATCCAAACGTTCACTTCACTTGGCTCCATCTTATGAGATGACCTACCGTCTTATG	1500
Db	1441	ATCCAAACGTTCACTTCACTTGGCTCCATCTTATGAGATGACCTACCGTCTTATG	1500
Qy	1501	TTCGTTTATCTCAACATGATGTGACCAATCACTTCGAGATGAGCAATCATGAGAT	1560
Db	1501	TTCGTTTATCTCAACATGATGTGACCAATCACTTCGAGATGAGCAATCATGAGAT	1560
Qy	1561	TGA 1563	
Db	1561	TGA 1563	

RESULT 2

ABK50985

ID ABK50985 standard; DNA; 1539 BP.

ABK50985;

24-SEP-2002 (first entry)

DNA encoding abscisic acid (ABA) hydroxylase CYP72A7.

ABA hydroxylase; abscisic acid hydroxylase; transgenic; plant; CYP72A7; gene; ds.

Arabidopsis thaliana.

Key Location/Qualifiers

FT 1..1539

FT /tag= a

FT /product= "ABA hydroxylase, CYP72A7"

PN MO500246377-A2.

PD 13-JUN-2002.

XX 06-DEC-2001; 2001MO-CA001756.

XX	07-DEC-2000; 2000US-0251518P.
XX	(CANADA) NAT RES COUNCIL CANADA.
XX	Krochko JE, Cutler AJ, Abrams SR;
XX	WPI; 2002-519663/55.
DR	P-PDB; AAU97096.
XX	New isolated and purified DNA that encodes protein having abscisic acid
PT	(ABA) hydroxylase activity, useful for altering catabolism of abscisic
XX	acid in plants.
XX	Claim 7; Page 104-105, 117pp; English.
CC	The invention relates to an isolated and purified DNA (I) that encodes a
XX	protein having abscisic acid (ABA) hydroxylase activity. (I) is useful
CC	for producing a transgenic plant which involves introducing (I) into a
CC	genome of the plant or its part, and carrying out plant growth and
CC	development. (I) is useful for modifying catabolism of ABA or ABA
CC	analogues in a plant. ABA hydroxylase clone pBE10-30-3 cDNA sequence is
CC	useful for identification of related sequences from other plant species.
CC	(I) is also useful for altering ABA and ABA analog catabolism in plants.
CC	The present sequence encodes abscisic acid (ABA) hydroxylase CYP72A7
SQ	Sequence 1539 BP; 452 A; 307 C; 371 G; 409 T; 0 U; 0 Other;
Query Match	20.4%; Score 318.8; DB 6; Length 1539;
Best Local Similarity	52.8%; Pred. No. 7.1e-85;
Matches 761; Conservative	0; Mismatches 672; Indels 9; Gaps 3
Dy	90 GTTATGCTGAGACCAAGAAGATTGAAGAAATTCTTTCTTAACAAGAAATTGAGTCCT 149
Qy	78 GCCTCGATTAAACCAAAGATCTTGAGAGTCTTTGAAAAGACGGGTCTTACCAGAAC 137
Db	150 TCCTTATCATTTCTTCATGGAATGTTAAAAACTGTTGGATGATGCTTTAAAGCTTC 209
Qy	138 TCCCTACACCTCCTCTGCTGGAGATATAAGAAAGATTGATATGATGATGGAACGAG 197
Db	210 TTCTCATATCCATGCCCTTCTCTCACAAATATCTCTAGAGTCTCTCTTTTACCANTCA 269
Qy	198 ATCTTAACCCATCAATGTAACGAGATGATACCCCAACGCCTCTCTCTGCTTAA 257
Db	270 CTGAGAAAAATCTACGCTGCTACATTTCTGGTTTTGTTGGTCCAATTCCGGTTAAC 329
Qy	258 GATGCTCAATTTCTACAGGAAGACCTTTCTATGATGATGATGACCTTCCACAGATTGT 317
Db	330 GGTAGCCGATCCGATTTGATACGAGAGATCTCTCTAAGCTGATGCTTACAGAAAGA 389
Qy	318 GATTAAGAAATCTGACAGATTAAGAAAGTCTTTTAATTAAGTCAAGACCTTTGAGAAAC 377
Db	390 TGAAGTCAACCTTTGGTTTAAACAATTGAAGGAGATGACTATTAGTCTCAAGGTGA 449
Qy	378 TTCTACATTCCTTTGATCAGATGTTGACAGS--TGGCTTGGCAAGTTACAGAGAGA 434
Db	450 AAAATGGGCTCATATCGAAAATATCATAGCCCTTAATTTCAATGAGAAATCTTAAGTT 509
Qy	435 TAATGGGCGAAGTCAAGAGATATCAACCCGGCTTTTCACTCGAAAATATCAAGA 494
Db	510 GCTTAGTACAGTTGTTGATAGAGTGTGATGATATGATGATTAATGATCCGATAAGTT 569
Qy	495 CATGATCCCGGCTTTTACCATTTGTTGACGAGAGGTGTCTGTCAATGGAGAAAGTATT 554
Db	570 ATCAAGAAAAG--GTGAAGTTGAGGTAGATGTCTAGAGGCTTTAGATTTTGACGA 626
Qy	555 TAAAGATTAAGAAAGTCCCTCTTGAAGTGAATGTTGGCTTGGTTTGAATATGACTGC 614
Db	627 AGATGTTAATAGTAAGAACAGTTTGGAAAGTATGATGATGATGATGATGATGATGATG 686
Qy	615 GGATGTCATCTGCACATCTGTTTGGAAAGTATGATTAAGAAAGGAGAGAAATATTCA 674
Db	687 ACTTCAAGTCAACAAATGCTTTTGTGCTGAGAGCTTTCAAAAAGTCTTCAATCTCG 746

D	b		675	ACTACAAAGGGGAATTGCGTGAAGCTTAATGCACAAGGCTTTAAGAATCTTACATCCCTGG	734
O	y		747	CATATGATTTTTTCCGACAAAGAGGAATTTGAAGTCGGAAGTTAGACAAAGAGATTAAG	806
D	b		735	ATCGAGGTTTACCACAAAGAGCAATTAAGAGTAGAAGAAATGATAGAGAGTGAGA	794
O	y		807	GAAATCGTTGTGAAAGCTGATAGACGGCGGAGACAAAACGCTATATGATGGAAGGGGA	866
D	b		795	CGTAATATTGAGAGGTAATTGAGCAAACGGAGAAAGCAGAGAAAGCTGGAGAACG	854
O	y		867	AGCATGTAGAGACCGCGCGGAGAGATTTGTGGGATTAATGA--TTCAAGCAAGANA	923
D	b		855	AATCATGATTTTGTGGGGAATTGCTGGAATCCAATTCAAGAGATCTCAAGAAATGG	914
O	y		924	TGTGACGGTTCAGACATTTGTGAGAGGTGTAAAGCTTTTCTTGGCCGGAAACAGC	983
D	b		915	AATGACCGTAGAAGATGTGATGAAGAGTGCAAGTTGTTTATTGCGGGACAAAGAC	974
O	y		984	AACCTCTAATCTGCGAGCGTGAAGACATCTTGATCCATGCAACCCGGAGTGGACGAC	1043
D	b		975	AACCTTAGTACTTTTGTGTGTGACATAAGTTTTATTAAGCATACCAAGACCTGGCAAGC	1033
O	y		1044	CAAGACCGTAGTAGAGTCTCTCAGAGGTCTCGGCTCACTGATGTCCCTTACCAAGACCA	1103
D	b		1035	TGCGGACGAGAGGAAGATGATGCAAGTACTCGGTGAATTAATTAACCTGATATGAGATC	1094
O	y		1104	TGTGCTAAGCTTAAAGSTTGAATGATCTTGAACGAGCTTTTAAAGCTTATGCCAC	1163
D	b		1095	CCTTAACAACCTTAAGCATGATATGATCTTCAATGAGTTTGAAGCTATACCTTCC	1155
O	y		1164	AATAGTAGTACGATTCGACGCGCTAATCGATGGAAGCTTAGAGGCTACAAAAATCC	1223
D	b		1155	GGTGGCTACGTTAAACAGATGTCAACAAAGAAATGAAGCTCGGAGAGTTGACCTTCC	1214
O	y		1224	ATGTGCGACGAGCTTCAATCCCAATCATAGCGGTCCATCATGACCAAGCCATTGGGG	1283
D	b		1215	AGCTGGAATTCAAATTAATCTTAACAATATTTCTTGTCAGCGTGAACCGAGCTTTGGGG	1274
O	y		1284	TATGACGTGAACGAATTCAAATCCAGCTCGGTTTGGGATGAGAGTCCGCGTGTGCCAA	1343
D	b		1275	CGATGATGACGAGATTTTAACCCGAGCGGTTCAAGAGAGGGCTCTCAAAAGCAACAA	1333
O	y		1344	ACACCCCGTGGCTTCAACCGTTTGGCTCGAGGTCTGTAATGATGATGGATGGTCAAGATCT	1403
D	b		1335	GAAACGAGTCTTCTTCTTCCCTTGGATGGGACCTAGATTTGCATCGGTCAAGATTT	1394
O	y		1404	TGCTAATCTTCAGGCAATTTGACACTCGCTGTATATGATCCAAACGCTTCACTTCACTT	1463
D	b		1395	TGCTATGTTGAGGCAAAAGATGGCAATGGCTTGATTTCTACAAAGTCTTCCCTTGAAGCT	1454
O	y		1464	GAGCTCTAATCTATACGATGACACTCAACGCTCTTAATGTTGCTTATCCCAACAGGTGC	1523
D	b		1455	CTCTCCTTCTTAATGTTCAAGCGCTCAACAGTCATACCACTGTCCCAATTTGGAGAC	1514
O	y		1524	AC 1525	
D	b		1515	TC 1516	
RESULT 3 ABL60758 standard; cDNA; 1584 BP.					
X	X		ABL60758;		
X	X		10-SEP-2002 (first entry)		
X	X		Maize cytochrome P450, CYP72A1 coding sequence.		
X	X		Maize, cytochrome P450; CYP72A1; CYP92A1; plant; gene therapy; antidote;		
X	X		gene; ss.		

Mon May 24 08:10:07 2004

us-09-992-901-1.rng

Page 4

OS Zea mays.
XX Key 1.1584 Location/Qualifiers
FH CDS //tag=a
FT //Product="CYP72A1"
XX
XX
XX US6380465-B1.
XX
XX 30-APR-2002.
XX
XX 11-JUL-1999; 99US-00351229.
XX
XX 12-JUL-1998; 98US-0092596P.
XX
XX (KENT) UNIV KENTUCKY RES FOUND.
XX
XX Barrett M;
XX
XX MPI: 2002-470227/50.
XX
XX P-PSDB; ABB08079.
XX
XX Nucleic acid molecules encoding a Zea mays cytochrome P450 CYP72A1,
XX useful for the production of plants resistant to heterocyclic,
XX sulfonylurea and substituted urea herbicides and organophosphate
XX insecticides e.g. Classic and Pursult.
XX
XX Claim 11; Col 9-12; 21pp; English.
XX
XX The invention relates to isolated nucleic acid molecules encoding Zea
XX mays (maize) cytochrome P450 designated CYP72A1 and CYP92A1. The
XX polypeptides can be expressed by standard recombinant methodology. The
XX nucleic acids may be used in genetic engineering protocols to transform
XX plants and other eukaryotes e.g. Yeast, maize (especially), soybean,
XX beer, tobacco, wheat, barley, poppy, rape, sunflower, alfalfa, sorghum,
XX rose, carnation, gerbera, carrot, tomato, lettuce, chicory, pepper, melon
XX and cabbage. They may be used in this way to confer protection against
XX heterocyclic herbicides, sulfonylurea herbicides, substituted urea
XX herbicides and/or organophosphate insecticides. The present sequence
XX represents the Z. mays CYP72A1 coding sequence
XX
SQ Sequence 1584 BP; 430 A; 389 C; 414 G; 351 T; 0 U; 0 Other;
Query Match 17.3%; Score 270.4; DB 6; Length 1584;
Best Local Similarity 51.5%; Pred. No. 2.7e-70;
Matches 680; Conservative 0; Mismatches 626; Indels 15; Gaps 2;
QY 228 CTCCTACATATTTCTCTAGAGTCTCTCTTTTACATCACTGAGAAAATCTACGG 287
DB 249 CTGCCACGACATCAACCCACGCGTGCAGCCCATGACATCAGACACCATCAAGAAATACGG 308
QY 288 TGCATCATTTCCGTTGGTTCGGTCCAACTTCCGGTTAAGCGTACCGCATCTGATT 347
DB 309 GAATCTATCTGCTTCACTGGTTCGCGCAACCAAGGATATATCTCTGACCCAGAGTT 368
QY 348 GATCAGAGAGATCTTCTCTAAGTCTGAGTTCTACGAGAGATGAAGTCACTCCCTTGGT 407
DB 369 AGTCAAGAGGCTGCTCTATATAAGTTGGCACTTTGGCAACCAAGAGATGACCGCAT 428
QY 408 TAAACACTTGAAGCGGATGAGTACTAGTCTCAAGGTAAGAAAATGGGCTCATCATCG 467
DB 429 TGGAGGTTGCTAGCGAAGCGGCTTGAATCATGATGTAAGAAAATGGGCAAGCAG 488
QY 468 AAAAATCATTACCCCTACTTTTCAATAGAGAAATCTTAAGTTGCTTGAACAGTTGGT 527
DB 489 GAGAAATCTTATCTGATTTCACTTCACTGAAAAAATGAAGGGATATGCTCCGATTTTC 548
QY 528 GAAGAGTGTGACTATATGATGATTAATGATGCTCCGATTAAGTTTCAAGAAAAGGTGAAGT 587
DB 549 TACCTGCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 608
QY 588 TGAGTGAATGCTTAATGATGATGATTTGATGATTTGATGATGATTTGATGATGATTTGATGAT 647

DB 609 TGAGATGATGCTGCTGCGCTGAGTTCAGAGATCTTACTGAGATGATTAATCAAGAACTGC 668
QY 648 TTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 707
DB 669 GTTTGGAGCACTATCAAGAGAGAGAGATTTTGGAGTCAAGAGAGAGAGAGAGAGAGAGAG 728
QY 708 TCTTGTGCTGAGCTTTTCAAAAAGTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 767
DB 729 AGCCCTCATCATCTGTTTCAAGCAATTTTATCCAGGCTATGTTCTTCCGCAACAA 788
QY 768 AGGAAATTTGAGTCTCGAAGTTTGAACAGAGATTAAGAAATGATGATGATGATGATGATGAT 827
DB 789 AAACACAGAGAGATGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 848
QY 828 AGAGCGCGGAGACAAAACGCTATAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879
DB 849 TGGAG 908
QY 880 ----CCGCGCGGAGAGATTTTGGATTAATGATTAATGATTAATGATTAATGATTAATGAT 935
DB 909 ATTACTGAGTAAACACAGAGCAATCAATGAGAAATGCAAGCTGGGATTTCAACAG 968
QY 936 GGCATTTGAGAGAGATGATTAAGCTTTTCTTCCGCGGAGAAACAGCAACTTCTATCT 995
DB 969 AGATGATGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1028
QY 996 GCTGACGTGAGAGACATCTTGTATCATGACCGGAGTGGACGAGCCAAACAGAGTA 1055
DB 1029 GCTTACTTGAACACTTATTTGCTAGATGACACCAAGATGGAGAGAGAGAGAGAGAGAG 1088
QY 1056 TGAAGTCTCAGAGTTCGCGGCTCAAGTATGCTTCAACAGAGACCATGCTTAAGCT 1115
DB 1089 AAGGTTTGGAGCCATTTGG---AAGAACACACAGATTAATGATTAATGATGATGATGATGAT 1145
QY 1116 TAAACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1175
DB 1146 CAGAGTGAATCAATGATTTCAATGAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGTCT 1205
QY 1176 GATTGAGCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1235
DB 1206 AACCAAGAGATCTTAAAGAAATGAGAGCTCGGAGATCAATTAATCTGAGAGAGTGA 1265
QY 1236 GCTTCAATCCCAATCAATGAGGCTCCATCAATGAGAGAGATGAGAGATGAGAGATGAGAG 1295
DB 1266 GCTTCTTCTGCGGCTCATCTTCAATCAATGATGATGATGATGATGATGATGATGATGAT 1325
QY 1296 GCAATTCATCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1355
DB 1326 CGAGTTCAACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1385
QY 1356 CTTTCAATCGTTTGGCTCGAGTTCATGATGATGATGATGATGATGATGATGATGATGAT 1415
DB 1386 TTTCTTCTGCTGAG 1445
QY 1416 GGCCAATTTGACATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1475
DB 1446 AAGCAAGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1505
QY 1476 TCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1535
DB 1506 CAGCAG 1565
QY 1536 C 1536
DB 1566 C 1566
RESULT 4
ABK50986
ID ABK50986 standard; DNA; 1539 BP.
XX AC ABK50986;
XX

DT 24-SEP-2002 (first entry)
 XX DNA encoding abscisic acid (ABA) hydroxylase CYP72A14.
 DE ABA hydroxylase; abscisic acid hydroxylase; transgenic; plant; CYP72A14;
 KW gene; ds.
 XX Arabidopsis thaliana.
 OS
 PH Key Location/Qualifiers
 FT CDS 1..1539
 FT /tag= a
 FT /product= "ABA hydroxylase, CYP72A14"
 XX
 XX W0200246377-A2.
 PD 13-JUN-2002.
 XX 06-DEC-2001; 2001MO-CA001756.
 XX 07-DEC-2000; 2000US-0251518P.
 XX (CANA) NAT RES COUNCIL CANADA.
 PA Krochko JE, Cutler AJ, Abrams SR;
 PI WPI; 2002-519663/55.
 DR P-PSDB; AAU97097.
 XX
 PT New isolated and purified DNA that encodes protein having abscisic acid
 PT (ABA) hydroxylase activity, useful for altering catabolism of abscisic
 PT acid in plants.
 XX
 PS Claim 13; Page 104-105; 117pp; English.
 XX
 CC The invention relates to an isolated and purified DNA (I) that encodes a
 CC protein having abscisic acid (ABA) hydroxylase activity. (I) is useful
 CC for producing a transgenic plant which involves introducing (I) into a
 CC genome of the plant or its part, and carrying out plant growth and
 CC development. (I) is useful for modifying catabolism of ABA or ABA
 CC analogues in a plant. ABA hydroxylase clone pB10-30-3 cDNA sequence is
 CC useful for identification of related sequences from other plant species.
 CC (I) is also useful for altering ABA and ABA analog catabolism in plants.
 CC The present sequence encodes abscisic acid (ABA) hydroxylase CYP72A14
 XX
 SQ Sequence 1539 BP; 452 A; 330 C; 369 G; 388 T; 0 U; 0 Other;
 Query Match 17.2%; Score 269.2; DB 6; Length 1539;
 Best Local Similarity 5.1%; Pred. No. 6.2e-70;
 Matches 737; Conservative 0; Mismatches 693; Indels 12; Gaps 4;
 QY 90 GTATGTGTGAGACCAAGAAAGATTGAAGAACTTCTTAACCAAGAAATCGAGTCC 149
 DB 81 GGTGTGTGTGACCAAGAAAGATTGAAGAACTTCTTAACCAAGAAATCGAGTCC 140
 QY 150 TCTTATCATTTTCTCATCGGAATGTTAAGAACTTGTGAAGTATGCTTAAGCTTC 209
 DB 141 TCTTACACGCTTAATCGCGCATTTTAAAGAAAGATGATGATGATGATGATGATG 200
 QY 210 TTTTATCATTTTCTCATCGGAATGTTAAGAACTTGTGAAGTATGCTTAAGCTTC 269
 DB 201 ATCCAAACCCATCAACCAAGATGATGATGATGATGATGATGATGATGATGATGATG 260
 QY 270 CTGAGAAATCTAAGGCTGCTCATTTCTGTTGGTGTGCTCACTTCCGCTTAAC 329
 DB 261 AATGCTCAAGCTCTGAGAGAGATTAATTAATGATGATGATGATGATGATGATGATG 320
 QY 330 GGTAGCCATCTGATTTGATGAGAGATTTCTTAAGTCTGAAGTTCTGAAGAGAA 389
 DB 321 CATATGATCTCTGAGCAATCAAGAAAGTTCACAAAGTCTATGATCTTCCAGAGGC 380
 QY 390 TGAAGCTACCTTGTGTTAAACAATTGAAGGCGATGAGTACTTATGCTCAAGAGTGA 449

DB 381 GCATACGTTTCTTTAAGCAAAATTAATCT---AGGCAAGGAGATCTGTTAGTATGAGCGA 437
 QY 450 AAAATGGGCTCATCAACCAAAATTAATAGCTTCTTCAATGAGAAATCTTAAGTT 509
 DB 438 TAAATGGGCGCAACCCGAAATATCAATCCGCTTTCCACCTTGAAGAGTCAAGAA 497
 QY 510 GCTTGTACCACTGTTGTTGAAGAGTGAATGATGATGATGATGATGATGATGATGATG 569
 DB 498 TATGTATGATGTTTCCACGAAGAGTGAAGAGTGTGTTGTTGTTGTTGTTGTTGTTG 557
 QY 570 ATCAGAAACCGTGA---AGTTGAGTATGATGATGATGATGATGATGATGATGATGATG 626
 DB 558 CTGAGTAAAGGATCTCATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 617
 QY 627 AGATGTTATTAATGAACAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 686
 DB 618 AGATGATGATCTCCGATCTGCTTTGTTGATGAGCTAAGAGAGAGCAAGAGATATTGA 677
 QY 687 ACTTCAAGCTCAAAATGCTTCTTGTGCTGAAGCTTTTCAAAAGTCTTCAATTCCTGG 746
 DB 678 ACTCAGAGAGAACTTAGACAGCTGATGAGCTTTTCAAGAAATTTTATTCCTGG 737
 QY 747 CTATAGATTTTTCGACAGAGGAAATTTGAAGTCTGGAAGTTAGACAGAGATTAAG 806
 DB 738 ATATATTATCTCCCAACAAAGGTTAATGAAGATGAAGAAACAGACGCCAGAGAAATCCA 797
 QY 807 GAAAGCTGTTGGAAGCTGATAGAGCGGCGAGACAAACGCTATAGATGAAGAGGGA 866
 DB 798 AGATATATCAAGAGGATCATTAACAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 857
 QY 867 AGATATGAAGAGCGCGCGAGAGAGATTTGTTGATTAATGATTAAGTCAAGAGAGAG- 925
 DB 858 AAGCAGAGATTTGCTAGTATCTTCTGAATCAAACTTGGCGCAAGAGAGAGATGG 917
 QY 926 --TGACGTTTCAGACATTTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 983
 DB 918 AATGATATCCAGAGATATATGAAGATGACAGTGTCTTCTTATTTGGCCGCGCAAGAGAG 977
 QY 984 AACTTGAATCTGCTGACGTGACGACATCTTGTATCATGACACCGAGATGAGAGG 1043
 DB 978 AATATAGATCTTCTGTTGACATGATGATGATGATGATGATGATGATGATGATGATG 1037
 QY 1044 CAAAGCAGTATGAGTCTTCAAGGCTGCGGCTGACGTATGATGATGATGATGATGATG 1103
 DB 1038 TCGTGACAGAGAGAGTGAAGCAATTTTGGCATTAACAACTGATTAAGAGGCT 1097
 QY 1104 TGTGTTAAGCTTAACAGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1163
 DB 1098 TAAAC---CAACTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATG 1154
 QY 1164 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1223
 DB 1155 TGTATGTCAGAGTACCGGACATTCACAAAGAGATGAAGCTGAGATCTGACTTAC 1214
 QY 1224 ATGTGACAGAGATCTTAATCCATCAATGAGGCTCATGATGACCAAGCATTTGGGG 1283
 DB 1215 AGCGGATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1274
 QY 1284 TAAATGAGTGAAGCAATTAATCAAGTGTGTTGGAGTGAAGTGAAGTGAAGTGAAGTGA 1343
 DB 1275 AAAGATGAGAGGAGATTCAGAGCTGAGATTAAGATTAAGAGAGGCTTCAAAAGCAACA 1334
 QY 1344 AAGACCGGTTGGCTTCAACGTTTGGCCGAGGATGATGATGATGATGATGATGATGATG 1403
 DB 1335 GAACCAAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1394
 QY 1404 TGCTATATCTTCAAGCCAAATTAAGTCACTGCTGTATGATGATGATGATGATGATGATG 1463
 DB 1395 TAAATGCTTGAAGCAAAATGAGCAATGATGATGATGATGATGATGATGATGATGATG 1454
 QY 1464 GGCCTCTATCTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1523
 DB 1455 TTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1514

Mon May 24 08:10:07 2004

us-09-992-901-1.rng

Page 8

```
DB      1324  GAAACGCAATTTCTCAAGCCACCATTCACCCAAAGCTCTCCCTCCGTTCTCAATCGGA 1383
QY      1378  GTTCGTACATGATGTCAGAAATCTGTCTAATCTTCAAGCCAAATGACACATCGCTGTA 1437
DB      1384  CCAAGAGCTTCATTCGCAAAACTTTCATGCTGAGCCAGACGCGTTCTCATATG 1443
QY      1438  ATGATCCACGCTTCACCTTCACTTGTCTCTACTTATCAGACATGACCTTACGTTCTT 1497
DB      1444  ATCCTTCAACAGTTCCAGCTTACCTCTCCCGAGTATACACACCGCAGTTGATCAC 1503
QY      1498  ATGTGCTTATCTCTCAACATGTCACCATCA 1531
DB      1504  TTGATCTCTTCCCGCATACGGGCTTACCGGTAA 1537

RESULT 7
AAC43009
ID      AAC43009 standard; DNA; 1712 BP.
XX
AC      AAC43009;
XX
DT      17-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana DNA fragment SEQ ID NO: 37690.
XX
KW      Hybridisation assay; genetic mapping; gene expression control;
KW      protein identification; signal transduction pathway; metabolic pathway;
KW      promoter; termination sequence; ss.
XX
OS      Arabidopsis thaliana.
XX
PN      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-00301439.
XX
PR      25-FEB-1999; 99US-0131825P.
PR      05-MAR-1999; 99US-0123180P.
PR      09-MAR-1999; 99US-0123548P.
PR      23-MAR-1999; 99US-0125788P.
PR      25-MAR-1999; 99US-0126264P.
PR      29-MAR-1999; 99US-0126785P.
PR      01-APR-1999; 99US-0127462P.
PR      06-APR-1999; 99US-0128234P.
PR      08-APR-1999; 99US-0128714P.
PR      16-APR-1999; 99US-0129845P.
PR      19-APR-1999; 99US-0130077P.
PR      21-APR-1999; 99US-0130449P.
PR      23-APR-1999; 99US-0130510P.
PR      23-APR-1999; 99US-0130891P.
PR      28-APR-1999; 99US-0131449P.
PR      30-APR-1999; 99US-0132048P.
PR      30-APR-1999; 99US-0132407P.
PR      04-MAY-1999; 99US-0132484P.
PR      05-MAY-1999; 99US-0132485P.
PR      06-MAY-1999; 99US-0132486P.
PR      07-MAY-1999; 99US-0132487P.
PR      11-MAY-1999; 99US-0132863P.
PR      14-MAY-1999; 99US-0134218P.
PR      14-MAY-1999; 99US-0134219P.
PR      14-MAY-1999; 99US-0134221P.
PR      14-MAY-1999; 99US-0134378P.
PR      18-MAY-1999; 99US-0134768P.
PR      19-MAY-1999; 99US-0135124P.
PR      20-MAY-1999; 99US-0135124P.
PR      21-MAY-1999; 99US-0135353P.
PR      24-MAY-1999; 99US-0135629P.
PR      25-MAY-1999; 99US-0136021P.
PR      27-MAY-1999; 99US-0136392P.
PR      28-MAY-1999; 99US-0136782P.
PR      01-JUN-1999; 99US-0137222P.
PR      03-JUN-1999; 99US-0137528P.
PR      04-JUN-1999; 99US-0137502P.
PR      07-JUN-1999; 99US-0137724P.
PR      08-JUN-1999; 99US-0138094P.
PR      10-JUN-1999; 99US-0138540P.
PR      10-JUN-1999; 99US-0138847P.
PR      14-JUN-1999; 99US-0139119P.
PR      16-JUN-1999; 99US-0139452P.
PR      16-JUN-1999; 99US-0139453P.
PR      17-JUN-1999; 99US-0139492P.
PR      18-JUN-1999; 99US-0139454P.
PR      18-JUN-1999; 99US-0139455P.
PR      18-JUN-1999; 99US-0139456P.
PR      18-JUN-1999; 99US-0139457P.
PR      18-JUN-1999; 99US-0139458P.
PR      18-JUN-1999; 99US-0139459P.
PR      18-JUN-1999; 99US-0139460P.
PR      18-JUN-1999; 99US-0139461P.
PR      18-JUN-1999; 99US-0139462P.
PR      18-JUN-1999; 99US-0139463P.
PR      18-JUN-1999; 99US-0139750P.
PR      18-JUN-1999; 99US-0139763P.
PR      21-JUN-1999; 99US-0139817P.
PR      22-JUN-1999; 99US-0139899P.
PR      23-JUN-1999; 99US-0140353P.
PR      23-JUN-1999; 99US-0140354P.
PR      24-JUN-1999; 99US-0140695P.
PR      28-JUN-1999; 99US-0140823P.
PR      29-JUN-1999; 99US-0140991P.
PR      30-JUN-1999; 99US-0141287P.
PR      01-JUL-1999; 99US-0141842P.
PR      01-JUL-1999; 99US-0142154P.
PR      02-JUL-1999; 99US-0142055P.
PR      06-JUL-1999; 99US-0142390P.
PR      08-JUL-1999; 99US-0142803P.
PR      09-JUL-1999; 99US-0142920P.
PR      12-JUL-1999; 99US-0142977P.
PR      13-JUL-1999; 99US-0143542P.
PR      14-JUL-1999; 99US-0143624P.
PR      15-JUL-1999; 99US-0144005P.
PR      16-JUL-1999; 99US-0144086P.
PR      16-JUL-1999; 99US-0144086P.
PR      19-JUL-1999; 99US-0144325P.
PR      19-JUL-1999; 99US-0144331P.
PR      19-JUL-1999; 99US-0144332P.
PR      19-JUL-1999; 99US-0144333P.
PR      19-JUL-1999; 99US-0144334P.
PR      19-JUL-1999; 99US-0144335P.
PR      19-JUL-1999; 99US-0144352P.
PR      20-JUL-1999; 99US-0144632P.
PR      20-JUL-1999; 99US-0144884P.
PR      21-JUL-1999; 99US-0144814P.
PR      21-JUL-1999; 99US-0145086P.
PR      21-JUL-1999; 99US-0145088P.
PR      22-JUL-1999; 99US-0145085P.
PR      22-JUL-1999; 99US-0145087P.
PR      22-JUL-1999; 99US-0145089P.
PR      22-JUL-1999; 99US-0145192P.
PR      23-JUL-1999; 99US-0145145P.
PR      23-JUL-1999; 99US-0145218P.
PR      23-JUL-1999; 99US-0145224P.
PR      26-JUL-1999; 99US-0145276P.
PR      27-JUL-1999; 99US-0145913P.
PR      27-JUL-1999; 99US-0145918P.
PR      27-JUL-1999; 99US-0145919P.
PR      28-JUL-1999; 99US-0145951P.
PR      02-AUG-1999; 99US-0146386P.
PR      02-AUG-1999; 99US-0146389P.
PR      02-AUG-1999; 99US-0146389P.
PR      03-AUG-1999; 99US-0147038P.
PR      04-AUG-1999; 99US-0147204P.
PR      04-AUG-1999; 99US-0147302P.
PR      05-AUG-1999; 99US-0147192P.
```

PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 23-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 30-AUG-1999; 99US-0151080P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155559P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159283P.
 PR 13-OCT-1999; 99US-0159284P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161950P.
 PR 28-OCT-1999; 99US-0161952P.
 PR 28-OCT-1999; 99US-0161953P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match

15.4%; Score 241.2; DB 3; Length 1712;

Best Local Similarity 49.2%; Pred. No. 1,8e-61;
 Matches 754; Conservative 0; Mismatches 768; Indels 12; Gaps 4;

QY 10 GAAAGTACAGCTGGTTCATTCCAAAGGTTCTTGTCCTCTGTAATCTTAGTCTGTA 69
 DB 4 GAAGCTTAAAGACAACTAACTCTTAACAGTCTCTTCTCTCTCTCTCTCTCTCTCAAG 63
 QY 70 ATAGTGAAGGTATGCTCTGTTATGTTAGTGAGACCAAGAAATGAAAGACATTTCTCT 129
 DB 64 ATATGAAAGCTTGTGATCAAG 123
 QY 130 AAACAGGAATTCAGAGTCT 189
 DB 124 AAACAGGAATTCAGAGTCT 183
 QY 150 GAAATGATGCTTAAAGCT--TCTTCTCATCTTATGCTCTCTCTCTCTCTCTCTCTCT 246
 DB 184 AAGATGAAGAAAGAGCTGACCTTGTGTTCTGATCCCAATTCATGATATTTCTCCT 243
 QY 247 AGAGTCT 306
 DB 244 CGTGTCT 303
 QY 307 TTGGTCCAACTTCCGTTTACGGTACCGATCTCTGATTTGATCAGAGATCTTCTC- 365
 DB 304 ACTGGAACAAACCACTATATATCATCTCAATCATGATGCTTACGAAACAGGCTTGTCTG 363
 QY 366 --TACGCTGAGTCTACGAAAGATTAAGCTCACCTTGTGTTAAACAATTGAGGC 423
 DB 364 AGCAAGTTCGTTCTACTATTATACCGTGAAGACCTGAGGCTTCTCATCTTTCTGCT 423
 QY 424 GATGACTACTTACTCTCAAGGTGAAGAAATGGGCTCATCTCAAAAACTTACCTT 483
 DB 424 AAAGACTTTCCTTTATCCAAAGGTGATGATGATGATGATGATGATGATGATGATGATGAT 483
 QY 484 ACTTTGATATGAGATCTTATGTTGCTTGTACCGAGTGTGTGAAGGTGTGATCT 543
 DB 484 GCTTTCCTCAGATCGGCTGACCGATGACCAACCGATGAGATGATGATGATGATGATGATGAT 543
 QY 544 ATGTGATTAATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 544 ATATTGAGAGTGAAGAAACAGAGAAATGGAAGTGTGATCAAGTATGATGATGATGATGATGAT 603
 QY 601 TATGATGCTTTCGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 604 AGCAAGATGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663
 QY 661 TATGAGATGCTGAGCAGTTTTCGATCTCAAGCTCAACAAATGCTTTTGTGCTGAA 720
 DB 664 TATGCGGAAGGATCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 723
 QY 721 GCTTTTCAAAAGCTTCATCTCTGCTATGATTTTTCGACAAGGGAATTTGAG 780
 DB 724 TCTCTCACTAACGTTTCATCT 783
 QY 781 TCTGGAAGTTCAGCAAGAGATTAAGAAATGCTGTTGTTGAAGCTATGAA--GCGGCGG 837
 DB 784 CTATGGAGGCTCCATAGAAAGTGAAGATCATCATCAAAAGATCATATGATTTCAAGGCTA 843
 QY 848 AGACAAAGCTATAGATGAGAGAGGGAAGATGTAAGAACCGCGCGGCAAGGATTTG 897
 DB 844 AATATCAAAATGTAAGCTTATGCTATGAGGAGATCTTCTAGGCTCATGTTGATCTCT 903
 QY 888 TTGGATTAATGATTCAGGCAAGAAATGAGCGTTCAAGACATTTGTGAGAGATGTA 957
 DB 904 GCGAATCTAAGATGATGAGAGAAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 963
 QY 958 AGCTTTTCT 1017
 DB 964 AATTTCTATGAGAGGCAAGATCTATCTCTATTTCTATGATGATGATGATGATGATGATGATGAT 1023
 QY 1018 CTATCCATGACCCGAGGTGAGGCAAGGCAAGGATGATGAGTCTCTAGGCTTCTGCGGC 1077

[illegible]

PR	19-APR-1999	99US-0130077P
PR	21-APR-1999	99US-0130449P
PR	23-APR-1999	99US-0130581P
PR	23-APR-1999	99US-0130891P
PR	28-APR-1999	99US-0131445P
PR	30-APR-1999	99US-0133046P
PR	30-APR-1999	99US-0133407P
PR	04-MAY-1999	99US-0133484P
PR	05-MAY-1999	99US-0133485P
PR	06-MAY-1999	99US-0133466P
PR	07-MAY-1999	99US-0133752P
PR	11-MAY-1999	99US-0133750P
PR	14-MAY-1999	99US-0133856P
PR	14-MAY-1999	99US-0134221P
PR	14-MAY-1999	99US-0134370P
PR	18-MAY-1999	99US-0133766P
PR	19-MAY-1999	99US-0133914P
PR	20-MAY-1999	99US-0135112P
PR	21-MAY-1999	99US-0135355P
PR	25-MAY-1999	99US-0136021P
PR	27-MAY-1999	99US-0136392P
PR	28-MAY-1999	99US-0136782P
PR	01-JUN-1999	99US-0131222P
PR	03-JUN-1999	99US-0137528P
PR	04-JUN-1999	99US-0137500P
PR	07-JUN-1999	99US-0138059P
PR	08-JUN-1999	99US-0138064P
PR	10-JUN-1999	99US-0138840P
PR	10-JUN-1999	99US-0139113P
PR	14-JUN-1999	99US-0133452P
PR	16-JUN-1999	99US-0133453P
PR	17-JUN-1999	99US-0133492P
PR	18-JUN-1999	99US-0133455P
PR	18-JUN-1999	99US-0133445P
PR	18-JUN-1999	99US-0133456P
PR	18-JUN-1999	99US-0133457P
PR	18-JUN-1999	99US-0133458P
PR	18-JUN-1999	99US-0133459P
PR	18-JUN-1999	99US-0133460P
PR	18-JUN-1999	99US-0133461P
PR	18-JUN-1999	99US-0133462P
PR	18-JUN-1999	99US-0133463P
PR	18-JUN-1999	99US-0133750P
PR	18-JUN-1999	99US-0133751P
PR	18-JUN-1999	99US-0133817P
PR	21-JUN-1999	99US-0133869P
PR	22-JUN-1999	99US-0140353P
PR	23-JUN-1999	99US-0140354P
PR	23-JUN-1999	99US-0140695P
PR	24-JUN-1999	99US-0140823P
PR	26-JUN-1999	99US-0140931P
PR	28-JUN-1999	99US-0141287P
PR	30-JUN-1999	99US-0141849P
PR	01-JUL-1999	99US-0142154P
PR	02-JUL-1999	99US-0142055P
PR	02-JUL-1999	99US-0142390P
PR	06-JUL-1999	99US-0142803P
PR	08-JUL-1999	99US-0142920P
PR	09-JUL-1999	99US-0142977P
PR	12-JUL-1999	99US-0143542P
PR	13-JUL-1999	99US-0143624P
PR	15-JUL-1999	99US-0144005P
PR	16-JUL-1999	99US-0144038P
PR	16-JUL-1999	99US-0144086P
PR	19-JUL-1999	99US-0144331P
PR	19-JUL-1999	99US-0144332P
PR	19-JUL-1999	99US-0144334P

PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145152P.
PR 23-JUL-1999; 99US-0145155P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145319P.
PR 28-JUL-1999; 99US-0145351P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147088P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147282P.
PR 05-AUG-1999; 99US-0147280P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149175P.
PR 17-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149922P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151086P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157733P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159285P.
PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.98; Score 185.4; DB 3; Length 1695;
Best Local Similarity 49.08; Pred. No. 1.1e-44;
Matches 751; Conservative 0; Mismatches 691; Indels 91; Gaps 6;

QY 90 GTTATGCTGAGACAGAAAGATTGAGAACTTTCTTAACAAGAAATTCGAGTTC 149
DB 139 GCTTTGCTTAAACCAAGATGCTTGAAGTTCCTGAGAAAGAAAGTTCGCCGGAAC 198
QY 150 TCTTTATCATTTCTTCATGCGAAATGTTAAAGACTTTGGATATGCTTAAAGCTTC 209
DB 159 TCTTTACACGCCCTTTCGCTGATTTGAAGAAATTTTACATGCTGCGAGGACG 258
QY 210 TTTCTATGCTATGCTTCTCTCAAAATTTCTTCAAGATTCCTTTTACATCA 269
DB 259 ATCCAAACCATCACTTACAGATATATCAACCAAGATTCGCTTACCGTTGA 318
QY 270 CTGAGAAATTTCAAGTCTCAATTTCTGTTGTTCCGTTCACTTCCGTTAC 329
DB 319 AATGCTCAAGCTCATGAGGAGACTTCTTACATGTTGAGCCATCAACAATAC 378
QY 330 GGTAGCCGATCTGATTTGATGAGAGATCTTCTTAAGTCTGAGTTCTACGAGAGA 389
DB 379 CATATGATGCTGAGCAATCAAAAGATTTCAAAAGTTATGATTTCCAAAGGC 438
QY 390 TGAAGCTACCTTTGTTAAACACTTGAAGGCGATGACTACTTACTTAAGTGA 449
DB 439 ACATAGCTTCTCTGCGAGATTAAAT--AGCGCTGACCTGTAATGATGATGTA-495
QY 450 AAAATGGCTCATCATGAAATCAATGACCTTCTTATGAGAAATCTTAAGT 509
DB 456 TAAATGACAAAGACCGAAGATCAACCGCGTTTCCACCTTGAAGATCAAAA 555
QY 510 GCTTGATACAGTTGTGTAAGATGATGATATGATGATGATGATGATGATGAT 569
DB 556 TATGATACCTGCTTCCACAGAGCTGACGAGATGTTGCGATGAGGACAAATGAT 615
QY 570 ATCAGAAAGCGTGA---CTTGAAGTATGCTATGATGATGATGATGATGATGAT 626
DB 616 GACGATTAACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGAT 675
QY 627 AGATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 686
DB 676 AGATGATATCTCCGATCTCATTTGACAGACTTCAAGAGGCGAGAGATTTTGA 735
QY 687 ACTTCAAGCTCAAAATGCTTTTGTGCTGAAGCTTTCAAAAGTCTTCAATTCCTGG 746
DB 736 GCTCAAGCGGAATTAACAGCTCATTAAGCTTTTGGAAAGCTATATCTCCGAG 795

QY 747 CTATAGATTTTCCGACAAAGGAAATTGAAGTCTCGAAGTTAGACAAAGATATAG 806
 DB 796 ATATGCTATCTCTCTCAAAAGGCTATATAGAG--ATGAAAGCGAGCGTATAGAAAT 852
 QY 807 GAATCGTTGTTGAAGCTGATAGACCGCGGAGCAAAAGCTATATATGAGAAAGGGA 866
 DB 853 CAATTTTACTGAGAGGATCTTTAACTAAAGGCTTAAAGGAGGAGCTGGGAGACA 912
 QY 867 AGAATGTAAGAGCGCGGCGGAGAGATTTGTTGAGATTATGATTCAGCAAAAGAT-- 924
 DB 913 CCAAGTGAAGATTTGCTGGGATATCTTCTGATTCGATTTGGCAAAACAAAGAAAT 972
 QY 925 ---GTGACGGTTACAGGACATTTGAGAGAGTGTAAAGCTTTTCTCCCGGAAACAG 981
 DB 973 GGAATGATGACCGAGGAACTGATGAGAGATGAGATTTGTTATTTCTGCTGGCAAG 1032
 QY 982 ACAATCTTAACTCTGACGTGAGACACATCTTCTATCCATGACACCGGAGTGGAG 1041
 DB 1033 ACTACTACAGTACTCTGTTGTTGACAAATGTTGTTAAGCCACACCAAGATTGGAG 1092
 QY 1042 GCCAAAGCACTGATGAGTCTCAGGGTCTGGGC-----TCAGTATGTCCTTACC 1095
 DB 1093 GCTCGTGCAGAGAAAGATGAAAGCAATTTTGGGATTAAGAACTGATGACAGAGGC 1152
 QY 1096 AAGGACCATGTC----- 1107
 DB 1153 TTGAACCAAGCTCAAAAGTTGATTAAGCAACCAAAACCTTCAAGTACGTGATCT 1212
 QY 1108 -----GTAACTTAAACGTTGATGATGATCTTGAACGAT 1144
 DB 1213 TTGCTGATTAAGCTGAAAGATTTGATCTCTCAAGATGACGATGATATATATAG 1272
 QY 1145 CTTTAAAGTTGATCTCAACCAATATGATGATGATGATGATGATGATGATGATGATG 1204
 DB 1273 TCTTAAAGCTATATCTCTCAAGTATGATGATGATGATGATGATGATGATGATG 1332
 QY 1205 TAGAGGATGAAATCCCATGATGATGATGATGATGATGATGATGATGATGATGATG 1284
 DB 1333 TAGGTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1392
 QY 1265 ATGACCAAGCAATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1324
 DB 1393 GCGAAGGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1452
 QY 1325 GATGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1384
 DB 1453 GTCTCTCAAAAGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATG 1512
 QY 1385 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1444
 DB 1513 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1572
 QY 1445 AAGCTTCACTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1504
 DB 1573 GGAATTTCTCTTGAAGCTTCTCTTATGATGATGATGATGATGATGATGATGATG 1632
 QY 1505 TTTATCTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1537
 DB 1633 CTCACCAAGTGGTGTCTCTCTCACTTTCG 1665

RESULT 11

AAA9071
 ID AAA9071 standard; DNA; 588 BP.

AC AAA9071;

DT 18-JAN-2001 (first entry)

DE Arabidopsis thaliana basal promoter sequence SEQ ID NO:16.
 XX Arabidopsis thaliana; basal promoter; cytochrome P450; CYP72B1; plant;
 XX brassinosteroid signaling; brassinosteroid synthesis; brassinolide; ds.
 KW

OS Arabidopsis thaliana.
 XX
 PN WO200055302-A2.
 XX
 PD 21-SEP-2000.
 XX
 PE 16-MAR-2000; 2000WO-US006915.
 XX
 PR 16-MAR-1999; 99US-0124570P.
 PR 14-DEC-1999; 99US-0170931P.
 PR 20-DEC-1999; 99US-0172832P.
 XX
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 XX
 PI Neff MW, Chory J;
 DR WPL; 2000-638195/61.
 XX
 PT Transgenic plants having modulated brassinolide synthesis resulting in
 PT insect resistance, dwarfism and darker-green foliage compared with wild-
 PT type plants, have nucleic acid encoding BAS1 polypeptide in its genome.
 XX
 PS Claim 79; Fig 1C; 104pp; English.
 XX
 CC The present invention describes a genetically modified plant (1)
 CC comprising at least one exogenous nucleic acid sequence encoding a BAS1
 CC polypeptide, homologue or functional fragment, in its genome or at least
 CC one regulatory sequence that modulated expression of endogenous basal gene,
 CC homologue or functional fragment, and which is characterized as having
 CC modulated brassinolide activity or synthesis. The basal gene encodes a
 CC cytochrome P450 (CYP72B1), which has a role in brassinosteroid signaling
 CC or synthesis. Overexpression of the basal gene in plants causes a dark
 CC green, dwarf phenotype which mimics plants that have low levels of the
 CC plant hormone, brassinolide. Overexpression of the basal gene also
 CC increases resistance to insects in plants. The present sequence
 CC represents the Arabidopsis basal promoter sequence, which is used in the
 CC exemplification of the present invention
 XX
 SQ Sequence 588 BP; 193 A; 121 C; 99 G; 173 T; 0 U; 2 Other;
 Query Match 11.8%; Score 184; DB 3; Length 588;
 Best local similarity 99.0%; Pred. No. 1.7e-44;
 Matches 206; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
 QY 1 ATGAGGAAAGAAAGTACGCTGTTATTCGAAAGTCTTCTGCTGTAATCTTA 60
 DB 362 ATGAGGAAAGAAAGTACGCTGTTATTCGAAAGTCTTCTGCTGTAATCTTA 441
 QY 61 AGT-CTGTAAATAGGAAAGGTATGCTCTGTTATGATGAGACCAAGAAATGAGAA 119
 DB 442 AGTCTTGTAAATAGGAAAGGTATGCTCTGTTATGATGAGACCAAGAAATGAGAA 501
 QY 120 AATTTCTTAAACAAAGAAATGAGAGTCTCTCTTATCTTCTTATGAGAAATGTA 179
 DB 502 AATTTCTTAAACAAAGAAATGAGAGTCTCTCTTATCTTCTTATGAGAAATGTA 561
 QY 180 AGAATTTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 207
 DB 562 AGAATTTGTT-GAATGATGATGATGATGATGATGATGATGATGATGATGATG 588

RESULT 12

AEN98320/C
 ID AEN98320 standard; DNA; 1124 BP.

AC AEN98320;

DT 01-AUG-2002 (first entry)

DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 88.
 XX Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;
 KW

KW disease crop; thale cress; tolerance factor; insect; pathogen;
 KM nutrition; ds.
 XX Arabidopsis thaliana.
 XX US2002023281-A1.
 XX
 PD 21-FEB-2002.
 XX
 XX 26-JAN-2001; 2001US-00770445.
 XX
 PR 27-JAN-2000; 2000US-0178472P.
 XX
 XX (GORL/) GORLACH J.
 PA (ANY/) AN Y.
 PA (HAM/) HAMILTON C M.
 PA (PRIC/) PRICE J L.
 PA (RAIN/) RAINES T M.
 PA (YUY/) YU Y.
 PA (RAME/) RAMEKA J G.
 PA (PAGE/) PAGE A.
 PA (MATH/) MATHW A V.
 PA (LED/) LEDFORD B L.
 PA (WOES/) WOESSNER J P.
 PA (HAAS/) HAAS W D.
 PA (GARC/) GARCIA C A.
 PA (KRICK/) KRICKER M.
 PA (SLAT/) SLATER T.
 PA (DAVI/) DAVIS K R.
 PA (ALLE/) ALLEN K.
 PA (HOFF/) HOFFMAN N.
 PA (HURB/) HURBAN P.
 XX
 XX Gorlach J, An Y, Hamilton CM, Price JT, Raines TM, Yu Y,
 PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
 PI Hurban P;
 XX
 DR WPI; 2002-403163/43.
 XX
 PS Claim 1; SEQ ID NO 88; 49pp + Sequence Listing; English.
 CC
 CC The invention relates to an Arabidopsis thaliana nucleic acid (I)
 CC comprising a sequence capable of hybridizing under stringent conditions
 CC to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
 CC given in the specification or its fragment. A polypeptide (II) encoded by
 CC (I), a transgenic plant (III) comprising an exogenous nucleic acid, or a
 CC genetically modified cell (IV) comprising an exogenous nucleic acid, is
 CC useful for screening a candidate agent for its biological effect. (I) is
 CC useful in identifying homologous or related genes, in producing
 CC compositions that modulate the expression or function of its encoded
 CC protein, mapping functional regions of the protein and in studying
 CC associated physiological pathways. (I) is also useful for the genetic
 CC manipulation of cells, particularly plant cells. (I) is also useful in
 CC screening assays of various plant strains to determine the strains that
 CC are best capable of withstanding a particular disease or environmental
 CC stress. (II) and (III) are useful for screening of biologically active
 CC agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
 CC pathways. The screened agents are useful in improved methods of treating
 CC crops to prevent or treat disease. (II) are also useful in screening
 CC programs to identify agents that mimic or enhance the action of tolerance
 CC factors. Such agents are useful in improved methods of treating crops to
 CC enhance their tolerance to environmental stress. (I) is also useful for
 CC enhancing or inhibiting production of a biosynthetic product in a plant.
 CC (III) is useful for identifying other mediators that may induce
 CC expression of proteins of interest, for establishing the extent to which
 CC any specific insect and/or pathogen is responsible for damage to a
 CC particular plant, for identifying other mediators that enhance or induce

Query Match	9.8%	Score 153.6	DB 6	Length 1124
Best Local Similarity	54.0%	Pred. No. 3,4e-35		
Mismatch 337	Conservative 0	Mismatch 284	Indels 3	Gaps 1
Qy	AGCCAAAGATGTGACGGTTCAGACATTTGTGAGGAGGTGTAAGCTTTTCTTCGCG	973		
Db	AAGCAACGATGTAGTACGAGATCTGATGAGGAGGTGCAAGTGTCTATTTCGCG	608		
Qy	974 GGAAGAGCAACTTCTATCTGTGACGTGAGACGACATCTGTATCATGACACCGCG	1053		
Db	607 GGCAGAGCAACATCATGACTCTTGTGTTGGCAATGGTCTCTTTAATCAACACGAG	548		
Qy	1034 AGTGCAGGCCCCAAGACAGTGTAGGTCTCCAGGCTCTGCAGGCTCAGGTATGTCCTTA	1093		
Db	547 ATGTGCAGGCTGTGTACAGAAAGATGTAGCAAGTGTGTCG--GATTAAGACCTG	491		
Qy	1094 CCAAGGACCATCTGTTAAGCTTAAACGTGTAGTGTATGATCTTGAACGAGCTTTAAGGT	1153		
Db	490 ATGCAAGAGATCTCAACAGCTCTAAAGTTATGACGATATTAATGAGTCTTGAGC	431		
Qy	1154 TGTATCCACCAATAGTAGTCTACGATTGCAGCGCTTAATCGAGTGAAGCTAGAGGCT	1213		
Db	430 TATATCTTCATATCCTCATGCTGACGACAGCCATTCACAAAGATGAGGCTGTGAGATC	371		
Qy	1214 ACAAATCCCATGTGTGACGAGGCTTATATCCCATGATPACGGCTCATGATGACCAAG	1273		
Db	370 TGCACATACAGGCGCGCTTGATCAATCTACTCTTCTTAATCCAGCGACAGCG	311		
Qy	1274 CCAATTGGGGTATATGACGTGAACGAATTCATCCAGCTGGTTGGGATGAGTGCCTC	1353		
Db	310 AGCTGTGGGGAACATGACAGAGAGTCAAGCAAGAGATTCAAAGCGTCTCTTAA	251		
Qy	1334 GTGCTGCCAAGACCCGTTGGCTTCATACCGTTGGCTCGAGATTCGATCATGCAATG	1393		
Db	250 AGGCAACAAGAACCAAGCCCTCTTCTTCTTTGCGTGGGACCGAGGATTCGCAATG	191		
Qy	1394 GTGAGATTTGCTATCTTATGAGGCCAAATGACATCGCTGTATGATCCAGCTTCA	1453		
Db	190 GCGAGATTTTGCATTTGTTGAGGCAAGATGCAATGATGATCCAGAGATTTCT	131		
Qy	1454 CCTTTCACCTGGCTCCTACTTATAGACATGCAACCTACCGTCTTATGTTGGCTTATCCTC	1513		
Db	130 CCTTGAAGTTTCTCTTCTCATATGTTCAGCACTTACAGAGTTCACATTCACCCAC	71		
Qy	1514 AACATGTGCACCAATCACCTTCG	1537		
Db	70 AGTTCGAGTCTCTCTTATATGTC	47		
RESULT 13				
ID	ADA48675			
XX	ADA48675 standard; DNA; 1377 BP.			
XX	ADA48675;			
XX	20-NOV-2003 (first entry)			
XX	Rice gene conferring disease resistance in plants.			
XX	disease resistance; pathogen tolerance; plant pathogen; ds; gene; plant.			
XX				

XX	WC2003000906-A2.
XX	
XX	03-JAN-2003.
XX	
PD	
XX	
PF	21-JUN-2002; 2002WO-IB002453.
XX	
DR	22-JUN-2001; 2001US-0300112P.
XX	
PR	26-SEP-2001; 2001US-0352277P.
XX	
PR	22-MAR-2002; 2002US-0365353P.
XX	
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.
XX	
XX	Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T,
PI	Katagiri F, Kreps J, Provart N, Rieke D, Zhu T;
XX	WPI; 2003-184052/18.
DR	P-PDB; ADA48676.
XX	
PT	New polynucleotide comprising a plant nucleotide sequence having an open
PT	reading frame that encodes a polypeptide associated with disease
PT	resistance, useful for conferring resistance or tolerance to a plant
PT	pathogen.
XX	
PS	ClaIm 1; SEQ ID NO 745; 239PD; English.
XX	
XX	
CC	The invention relates to a novel isolated polynucleotide comprising a
CC	plant nucleotide sequence having an open reading frame that encodes a
CC	polypeptide associated with disease resistance or its fragment having
CC	substantially the same activity as the full-length polypeptide. The
CC	polynucleotide of the invention is useful for conferring resistance or
CC	tolerance to a plant pathogen. The present sequence represents a gene
CC	conferring disease resistance used in the invention.
XX	
XX	Sequence 1377 BP; 324 A; 358 C; 389 G; 306 T; 0 U; 0 Other;

	Query Match	9.1%	Score 141.8;	DB 8;	Length 1377;
	Best Local Similarity	54.1%;	Pred. No. 1,4e-31;		
	Matches 312;	Conservative 0;	Mismatches 262;	Indels 3;	Gaps 1;
QY	970	GCCGGGAAACAGACAACTTCTPATCTGTGACGTGACAGCAATCTTGCTATCATGAC	1029		
Db	796	GCATGGCAATGACACCACTACACTGTGCTACCTGGACCAATGTTCTGTGTGACAGCAC	855		
QY	1030	CCGGAATGGCAGGCCAAAGCAGTGAATAGTCTCGAGGTGCGGGCTCAGCTATATC	1089		
Db	856	CCAGATTGGCAGGAATAATTAAAGGAGCAATTCGAATGGAAGTGTGTACAC--AGTGT	912		
QY	1090	CCTACCAAGACCATGTGTTAAGCTTAAACGTTGATGATCTTGAACAGATCTTTA	1149		
Db	913	CCCACTGGTGAATGCTCAACAAGCTAAAGTGTCAACATGTTCTCTTAAGACCTTA	972		
QY	1150	AGGTTTATCCACCAATGTAGTACATTCGAGGCGCTAAATCCGATGTGAAGGTAGA	1209		
Db	973	AGGCTGTACAGCCTCTGTCTACTCATACGAGGAAGTTGACACTGATATTGAGCTCGGC	1032		
QY	1210	GGGTACAAATCCCATGTGGCAGGAGCTTTAATCCATCATACCGGTCACTATGAC	1269		
Db	1033	GGCATCAAAATCCTGAGGGGTGCATATTGACGATCCCGATCCGCAATTCACCGGTAC	1092		
QY	1270	CAAGCATTTTGGGGTATAGAGGTGAACGAATTCATCCAGCTGGTTTCGAGTGAAGTG	1329		
Db	1093	AAAGAACTGTGGGGGAGAGATGCGAATGATTAAAGCCGAGAGGTTTCAGAAATGGGGT	1152		
QY	1330	CGGCTGTGCGAAACACCCCGTTGGCTTCAACGTTTGGCTCGGAATTGTATCATATC	1389		
Db	1153	ACGAGGGCAGCAACACCCCAAGCACTACTCTCTTCACGACGAGCCGAGGTCAATGC	1212		
QY	1390	AATGGTCAAAATCTTGCTATATCTTACGAGCCAAATGACATCGCTGTATATGATCAATGC	1449		
Db	1213	ATTGGGCAAACTTTTGCAATGATGAGGCCAAGGCTGTATTTGCCATGATCTCTAAGAGG	1272		
QY	1450	TTCACTTTCACTTGGCTCTACTTATTCAGATGACCTACGCTCTTATGTGCTTAT	1509		

Db	1273	TTCCTTCACTTCTTCTCCGCCCAAGTACGTCATGCAGCCACCTGATGTGATCAAGCTGCGT	1332
Qy	1510	CCTCAACATGCTGACCCATACCTTCCGAGATTGA	1546
Db	1333	CCCAAGTACGGGCTCCCTATGATCTCAAGAGCTCA	1369
RESULT 14			
ID	AAC54724/c		
AC	AAC54724	standard; DNA; 456 BP.	
XX			
DT	18-OCT-2000	(first entry)	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 78857.		
XX			
XX	Hybridisation assay; genetic mapping; gene expression control;		
KM	protein identification; signal transduction pathway; metabolic pathway;		
KW	Promoter; termination sequence; ss.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-00301439.		
XX			
XX	25-FEB-1999; 99US-0121825P.		
FR	05-MAR-1999; 99US-0123180P.		
FR	09-MAR-1999; 99US-0135348P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
FR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	28-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
FR	30-APR-1999; 99US-0132407P.		
FR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	06-MAY-1999; 99US-0132487P.		
PR	07-MAY-1999; 99US-0132863P.		
PR	11-MAY-1999; 99US-0134256P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	14-MAY-1999; 99US-0134370P.		
PR	18-MAY-1999; 99US-0134768P.		
PR	19-MAY-1999; 99US-0134941P.		
PR	20-MAY-1999; 99US-0135124P.		
PR	21-MAY-1999; 99US-0135353P.		
PR	24-MAY-1999; 99US-0135629P.		
PR	25-MAY-1999; 99US-0136021P.		
PR	27-MAY-1999; 99US-0136392P.		
PR	28-MAY-1999; 99US-0136782P.		
PR	01-JUN-1999; 99US-0137222P.		
PR	03-JUN-1999; 99US-0137528P.		
PR	04-JUN-1999; 99US-0137502P.		
PR	07-JUN-1999; 99US-0137724P.		
PR	08-JUN-1999; 99US-0138094P.		
PR	10-JUN-1999; 99US-0138540P.		
PR	10-JUN-1999; 99US-0138847P.		
PR	14-JUN-1999; 99US-0139119P.		

Mon May 24 08:10:07 2004

us-09-992-901-1.rng

PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139753P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0143977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144804P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145244P.
PR 26-JUL-1999; 99US-0145244P.
PR 27-JUL-1999; 99US-0145213P.
PR 27-JUL-1999; 99US-0145186P.
PR 27-JUL-1999; 99US-0145197P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 06-AUG-1999; 99US-0147493P.
PR 08-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156536P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157665P.
PR 07-OCT-1999; 99US-0158032P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 14-OCT-1999; 99US-0159844P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161935P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 6.7%; Score 104.2; DB 3; Length 456;
Best Local Similarity 52.7%; Pred. No. 1.6e-20;
Matches 226; Conservative 0; Mismatches 203; Indels 0; Gaps 0;

QY 1109 TTAAGCTTAAACGTTGAGTATGATCTTGAACGAGTCTTTAAGGTGTATCCACCAATAG 1168
Db 439 TTTTGTCTTACGATGACAAAGATTATATATGAGTCTTGAAGCTTATATCTTCCAGTAA 380

Mon May 24 08:10:07 2004

us-09-992-901-1.rng

Page 18

```
QY      1169  TAGCTACGATTCGACGCGCTTAATCGATGTGAGACTAGAGGGTACAAAATCCATGTG 1228
Db      379  TCAGATGACAGAGGAGCATCTCAAAAGAGATTAACCTAGGTGATATGACATACAGGCG 320
QY      1229  GACGAGGCTTCTAATCCCATATGCGGCTCCATCATGACCAAGCCATTGGGGTATG 1288
Db      319  GATATCAAGTCCATATGCGCTTCTGCTAATCCAGCATTAAGCTTTGGGGCGATG 260
QY      1289  AGCTGAACGATTCATTCAGCTCGGTTGGCGATGAGATGGCGCGTCTGCCAAAC 1348
Db      259  ATCCAGCGGAGTTCAAGCCCGAGAGGTTCAAAGACGGATCGCAAGGCAACAAAGAAC 200
QY      1349  CCGTTGCTTCATACCGCTTTGCGCTCGAGTTGCTACATGCAATGCTCAATCTTGCTA 1408
Db      199  AGGTCGCTTCTTGCCTTTGCGGTGGGACCGAGGATTTGCATCGGTCAAGACTTCGCTC 140
QY      1409  TACTTCAGGCCAATTCAGCACTCGCTGTAAATGATCCAAAGCTTCACCTTTCAGTTGGCTC 1468
Db      139  TGTGGAGCAAGATGCTCTGGCTTTATTTCTTCAGAGATTCTTCCTTCAGACTTTCAC 80
QY      1469  CTACTTATCAGCATGACCTACCGTCCCTTATGCTTATGCTTATCTCAACATGGTCAACCA 1528
Db      79  CTTCCTATGTTCAATTCACCTTACAGATCTTCACTATATACATCCACAGTGGGGCTCATC 20
QY      1529  TCACCTTCC 1537
Db      19  TTATCTCTC 11

RESULT 15
AAC53115/c
ID AAC53115 standard; DNA: 458 BP.
XX
AC AAC53115;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 73436.
XX
KM Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PM EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0133180P.
PR 09-MAR-1999; 99US-0133588P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134266P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-01343768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137223P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 26-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144353P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
```

Mon May 24 08:10:07 2004

us-09-992-901-1.rng

Page 19

PR 23-JUL-1999; 99US-0145224P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145913P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0145986P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148312P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151348P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153707P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154799P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.

PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 6.6%; Score 102.6; DB 3; Length 458;
Best Local Similarity 52.4%; Pred. No. 4.ee-20;
Matches 225; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

QY 1109 TTAAGCTTAAACCTTGAATGATCTTGAACGAGTCTTAAGTTGATCCACCAATAG 1168
DB 441 TTTGTTCTTCAAGATGCAATGATTTATATGAAAGTCTGAGCTATATCTCCAGTAA 382
QY 1169 TAGCTACGATTCGACGCGCTAAATCGGATGAGCTAGGAGGGTACAAAATCCCATGTG 1228
DB 381 TCCAGATGACAGAGCGGCTGACAAAAGATTAAGCTAGATGATGACACTACAGGCT 322
QY 1229 GACGAGCTTCTATATCCCAATCATAGCGGTCCATCATGACCAAGCCATTGGGGTAATG 1288
DB 321 GTATTCAAGTCCATATGCTGTGTTGCTAAATCCACCGTGAOACTAAGCTTTGGGGCGATG 262
QY 1289 ACGTGAAGAAATTCATCCAGCTCGGTTTGGGATGAGTGGCGGTGTCGCAAAACC 1348
DB 261 ATGCGAGGAGTTCAAGCCGAGAGGTTCAAAAGAGGATCGCAAGGCAAAAGAAC 202
QY 1349 CCGTTGGCTTCAATCCGTTTGGCTCGAGTTCTATATGATGATTTGGTCAATCTTGCTA 1408
DB 201 AGGTCGCTTCGCTTCGCGGTGGGACCGAGATTTGATCGTCAAGTCTTC 142
QY 1409 TACTTCAGGCCAAATGACACTCGCTGTAATGATTCACAGCTTCCTTCACTTGCTC 1468
DB 141 TGTGGAAGCAAAAGTGGCTGTGCTTGTATCTTCAGAGATTCCTTGAGCTTCAC 82
QY 1469 CTACTTATCAGATCAGCTTACCGTCTTATGTTGCTTATCTTCACATGATGACCAA 1528
DB 81 CTTCCTATGATTCATTCACCTTACAGAGCTTCACTATACATCCACAGTGTGGGGCTATC 22
QY 1529 TCACTTGC 1537
DB 21 TTATCCTGC 13

Search completed: May 22, 2004, 17:15:54
Job time : 695 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2004, 16:11:26 ; Search time 4335 Seconds
(without alignments)

10766.918 Million cell updates/sec

Title: US-09-992-901-1
Perfect score: 1563
Sequence: 1 atgagagaagaagtagcag.....tgacacatcatgagattga 1563

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapekt 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estbda.*
2: em_estbnum.*
3: em_estlin.*
4: em_estlma.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hcc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hcc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rpd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	28.3	562	B29839	B29839 T17P14TFC T
2	345.8	22.1	727	B1926826	B1926826 EST546715
3	316.2	20.2	558	B28185	B28185 T4A14TFC TA
4	311.6	19.9	916	CK262340	CK262340 EST708418

Result	Score	Query Match	Length	ID	Description
5	307.8	19.7	614	B0891384	B0891384 P049901 P
6	287.2	18.4	1876	A109087	A109087 Zea mays
7	280.2	17.9	698	CG027129	CG027129 CHGA11TFR
8	277	17.7	807	CK262341	CK262341 EST708419
9	259.6	16.6	817	BH597310	BH597310 BOGT577R
10	256	16.4	493	B0297122	B0297122 sa038e02
11	244	15.6	303	B62694	B62694 T17P24TFC.1
12	242.4	15.5	834	CG112869	CG112869 PURK51TB
13	234	15.0	1567	A1104671	A1104671 Zea mays
14	232.8	14.9	825	AV417119	AV417119 AV417119
15	219.6	14.0	414	CG336779	CG336779 OGUA25TV
16	206.2	13.2	377	AU287860	AU287860 AU287860
17	205.4	13.1	511	BJ558923	BJ558923 BJ558923
18	201.4	12.9	675	CP432662	CP432662 N1T1.18 B
19	199.8	12.8	445	BM144017	BM144017 sa151d11
20	198.8	12.7	861	A107534	A107534 Zea mays
21	197.6	12.6	533	BG556624	BG556624 EM1.37.E1
22	196.8	12.6	660	BH878919	BH878919 hsb5b04.g
23	196.4	12.6	412	BE210430	BE210430 sc043f12.Y
24	195.2	12.5	474	BG135565	BG135565 EST468457
25	193.2	12.4	730	CG027138	CG027138 CHGA11TFR
26	192.6	12.3	622	AM738972	AM738972 EST39499
27	189.6	12.1	968	BG321478	BG321478 Ds01.07G1
28	188.2	12.0	534	CD055580	CD055580 H009K04S
29	186.6	11.9	439	BH723464	BH723464 BOMIR22TF
30	185.8	11.9	649	AM622085	AM622085 EST112883
31	185.6	11.9	774	CF672445	CF672445 RTCTWT1.63
32	184.6	11.8	817	CK253045	CK253045 EST736682
33	184	11.8	87	AV920529	AV920529 AV920529
34	183.4	11.7	751	BE705698	BE705698 SC02.01h0
35	182.4	11.7	831	BE705698	BE705698 FGAS01516
36	180.8	11.6	1055	CK162569	CK162569 WHE2812.C
37	180.6	11.6	682	BQ801263	BQ801263 SALK_0803
38	180	11.5	223	CC054934	CC054934 AV944488
39	179.8	11.5	690	AV944488	AV944488 PAC000000
40	177.2	11.3	347	BME00512	BME00512 EST545534
41	176.4	11.3	738	B1925645	B1925645 NF078A02L
42	176	11.3	670	BG452638	BG452638 PCS03751
43	175.8	11.2	606	CA901199	CA901199 CD482646
44	174.6	11.2	715	CD482646	CD482646 at01-50m
45	174.6	11.2	754	B5574891	B5574891 w_A2R T1

ALIGNMENTS

RESULT 1 B29839 562 bp DNA linear GSS 13-OCT-1997
LOCUS T17P14TFC TAMU Arabidopsis thaliana genomic clone T17P14, genomic
DEFINITION Survey sequence.

ACCESSION B29839
VERSION B29839.1 GI:2515805

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 562)
Rounsley,S.D., Kelley,J.M., Field,C.E., Craven,M.B., Adams,M.D. and
Venner,J.C.

TITLE Use of a BAC End Sequence Database To Identify Minimal Overlaps for
Arabidopsis Genomic Sequencing

JOURNAL Unpublished (1997)

COMMENT Other GSSs: T17P14TFC
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@igf.org

Seq primer: M13-21
 Class: BAC ends
 High quality sequence atop: 562.
 Location/Qualifiers

FEATURES

source

1..562
 /organism="Arabidopsis thaliana"
 /mol_type="Genomic DNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="T17P14"
 /sex="hermaphrodite"
 /clone_lib="TAMU"
 /note="Vector: BelobAC11; Site_1: HindIII; Site_2:
 HindIII; Produced by Rod Wang"

ORIGIN

Query Match 28.3%; Score 442; DB 28; Length 562;
 Best Local Similarity 100.0%; Pred. No. 4.6e-116; Indels 0; Gaps 0;
 Matches 442; Conservative 0; Mismatches 0;

1122 GTTGAATGATCTTGAACGAGCTTTAAGGTTGATCACCAGTATGATGATTCG 1181
 493 GTTGAATGATCTTGAACGAGCTTTAAGGTTGATCACCAGTATGATGATTCG 424
 1182 AGCGCTTAATCGATGCTGAAGCTAGAGGGTACAAAATCCATGTGGACGGAGCTTC 1241
 433 AGCGCTTAATCGATGCTGAAGCTAGAGGGTACAAAATCCATGTGGACGGAGCTTC 374
 1242 AATCCCAATCATAGGCGTCCATCATGACCAAGCCATTGGGGTATGACGTGACGAT 1301
 373 AATCCCAATCATAGGCGTCCATCATGACCAAGCCATTGGGGTATGACGTGACGAT 314
 1302 CAATCCAGCTCGGTTTGGCGATGAGTGGCGGCTGCTCCAAACACCCCGTTGGCTTCAT 1361
 313 CAATCCAGCTCGGTTTGGCGATGAGTGGCGGCTGCTCCAAACACCCCGTTGGCTTCAT 254
 1362 ACCGTTGGCGTGGAGTTCGATACGATGCTGATGATGCTGATGATGCTGATGATGCTGAT 1421
 253 ACCGTTGGCGTGGAGTTCGATACGATGCTGATGATGCTGATGATGCTGATGATGCTGAT 194
 1422 ATTGACACTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1481
 193 ATTGACACTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 134
 1482 TGACACTACCGTCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 1541
 133 TGACACTACCGTCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 74
 1542 ATTGACCAATCATGAGATGGA 1563
 73 ATTGACCAATCATGAGATGGA 52

RESULT 2 727 bp mRNA linear EST 18-OCT-2001
 B1926826 EST546715 tomato flower, buds 0-3 mm Lycopersicon esculentum cDNA
 LOCUS B1926826
 DEFINITION clone cTOA30H6 5' end, mRNA sequence.
 ACCESSION B1926826.1 GI:16235694
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE

AUTHORS

TITLES

JOURNALS

COMMENTS

Van der Hoeven, R.S., Bezzerides, J.L., Karamycheva, S.A., Tsai, J.,
 Utechtback, T., Van Aken, S., Romling, C.M., Nieman, W., Fraser, C.M.,
 Martin, G.B., Giovannoni, J.U. and Jansley, S.D.
 Generation of ESTs from tomato flower tissue, 0-3 mm buds (2001)
 Unpublished (2001)
 Contact: CUGI

Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics
 Institute

FEATURES

source

Seq primer: T3.
 Location/Qualifiers
 1..727
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultiivar="TA96"
 /db_xref="taxon:4081"
 /clone="cTOA30H6"
 /tissue_type="flower"
 /dev_stage="0-3mm buds"
 /clone_lib="tomato flower, buds 0-3 mm"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI, Site_2:
 XhoI; supplier: Cornell University; sequencing: The
 Institute for Genomic Research; flower buds and flowers
 were taken from greenhouse plants (4-8 wks old, TA96).
 They were immediately frozen in liquid nitrogen and then
 size-separated while remaining frozen."

ORIGIN

Query Match 22.1%; Score 345.8; DB 12; Length 727;
 Best Local Similarity 72.7%; Pred. No. 3.4e-88;
 Matches 475; Conservative 0; Mismatches 172; Indels 6; Gaps 2;

34 AAGGTCCTGTCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 93
 78 AAGATTAATGAGTAATTTGTTGCTTCAATTTATGATGATGATGATGATGATGATGATGATGATGAT 137
 94 TGGTGAGACCAAGAAAGATGGAAGACATTTCTTAACAAGAAATTCAGGCTCTGCT 153
 138 TGGTTAACAAGAAAGAAATTTGAAACATTTTGTCTTAACAAGATTTAGAGGTCCACCT 197
 154 TATCATTTTCTTATCGGAATGTTAAAGAACTTGTGGAATGATGCTTAAAGCTTCTCT 213
 198 TATTAATTTTCAATTTGGAATGCTAAAGAAATTTGTTAATGCTAAAGGCTTCTCT 257
 214 CATCTATGCTTCTCTCAATATTTCTTCAAGTTCCTCTTTTACATCACTGG 273
 258 CAATCTATGCTTCTTATCAACAAATTTCTACAAAGATTTCTTATCTATCAATCAATGG 317
 274 AGAAATCTACGCTGCTACATTTCTGTTTGTGCTGCTCAACTTTCCGTTAAGGTA 333
 318 AGAAATTTTATGAGCAACATTTTATGTTGTTGGGCAACACACCTTTGGCGTGG 377
 334 GCCGATCTGATTTGATGAGAGATCTCTCT---AAGCTGAGTCTTACGAAAGAT 390
 378 GCTGATCTGATCTTATTAAGAAATTTTCACTCAAAATCTGAATTTTATGAAAAAT 437
 391 GAAGCTCAACCTTTGTTAAACAATTGAAGCGATGACTTACTTACTCAAGGTGA 450
 438 GAAGCTCAATCATTTATTAACAATTGAAGAGATGCTTCTTACTTAAAGGTGA 497
 451 AATGGGCTCATCAAGAAATTTATTAAGCTTCTTCTTCAATGAGAAATCTTAAGTTG 510
 498 AATGGGCTCATCAAGAAATTTATTAAGCTTCTTCTTCAATGAGAAATCTTAAGTTG 557
 511 CTGTGACCAAGTGTGTTGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
 558 TTAGTACCAAGGCGGAGTGAAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 616
 571 TGAAGAAACGTTGAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
 617 --AAGAAATGAGAAATTTGAATGAAATTAACAGAGTGTTCAGCTTTACAGAAAC 674
 631 GTTATTAAGTAAGACGTTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 683
 675 ATGTAAGTCAACAGCTTTTGGCCATGTTATGAACAGAAAGCTATTTT 727

RESULT 3
B28185/c
LOCUS B28185 558 bp DNA linear GSS 13-OCT-1997
DEFINITION T4A14TFG T4M1 Arabidopsis thaliana genomic clone T4A14, genomic survey sequence.
ACCESSION B28185
VERSION B28185.1 GI:2514151
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 558)
AUTHORS Rounsley,S.D., Kelley,J.M., Field,C.E., Craven,M.B., Adams,M.D. and Venter,J.C.
TITLE Use of a PAC End Sequence Database To Identify Minimal Overlaps for Arabidopsis Genomic Sequencing
JOURNAL Unpublished (1997)
COMMENT Other GSSs: T4A14TRB
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@igr.org
Seq primer: M13-21
Class: BAC ends
High quality sequence stop: 558.
FEATURES
source
Location/Qualifiers
1..558
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T4A14"
/sex="hermaphrodite"
/clone_id="T4M1"
/note="Vector: Beloskici; Site_1: HindIII; Site_2: HindIII; Produced by Rod Wang"
ORIGIN
Query Match 20.2%; Score 316.2; DB 28; Length 558;
Best Local Similarity 99.1%; Pred.No.1.1e-79; Mismatches 3; Indels 0; Gaps 0;
Matches 318; Conservative 0
1243 ATCCCAATCATAGCGGTCCATCATGACCAACCAATTTGGGGTAAATGACGTGAACGAATTC 1302
Db ATCCCAATCATAGCGGTCCATCATGACCAACCAATTTGGGGTAAATGACGTGAACGAATTC 499
QY 1303 AATCCAGCTCGCTTTGGCGATGAGATGCCGGTGTCTGCGCAAAACCCCGTTGGTTCATA 1362
Db 498 ATTCCAGCTCGCTTTGGCGATGAGATGCCGGTGTCTGCGCAAAACCCCGTTGGTTCATA 439
QY 1363 CGCTTGGGCTCGAGATTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1422
Db 438 CGCTTGGGCTCGAGATTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 379
QY 1423 TTGAACCTGCTGTATATGATATCAACGCTTACCTTCTGCTGCTTCTTCTTCTTCTTCTTCT 1482
Db 378 TTGAACCTGCTGTATATGATATCAACGCTTACCTTCTGCTGCTTCTTCTTCTTCTTCTTCT 319
QY 1483 GGACCTACCGTCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1542
Db 318 GGACCTACCGTCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 259
QY 1543 TTGACCAATCATGAGGATTGA 1563
Db 258 TTGACCAATCATGAGGATTGA 238

RESULT 4
CK262340
LOCUS CK262340 916 bp mRNA linear EST 12-DEC-2003
DEFINITION EST708418 potato abiotic stress cDNA library Solanum tuberosum cDNA clone POABF61 5' end, mRNA sequence.
ACCESSION CK262340
VERSION CK262340.1 GI:39819318
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 916)
AUTHORS Buehl,C.R., Hart,A., Zismann,V., Karanycheva,S.A. and Baker,B.
TITLE Generation of ESTs from abiotic stressed potato tissue
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST708419
Contact: Robin Buehl
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from TIGR via potatoc@igr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
source
Location/Qualifiers
1..916
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POABF61"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/clone_id="potato abiotic stress cDNA library"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."
ORIGIN
Query Match 19.9%; Score 311.6; DB 14; Length 916;
Best Local Similarity 63.4%; Pred.No.3e-78;
Matches 511; Conservative 0; Mismatches 289; Indels 6; Gaps 2;
37 GTTCTGTTGCTGTGATATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 96
Db 114 GTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 173
QY 97 TGGAGACCAAGAAAGATTGAAGACATTCTCTAAACAGAGATTGAGGCTCTCTAT 156
Db 174 TGGAGACCAAAATATATTCACCAACCTTCAAGAGACAGATCACAGGCTCTTAATAT 233
QY 157 CATTTCTTCATGGAATGTTAAAGAACTTGTGAATGATGCTTAAAGCTTCTTCAT 216

Db	234	CAATTTCTCTTGGTAACCTGAAGA---ATTGCAAGTTTCACTACTCCATCTTGAGCTT	290
Qy	217	CCTATGCGCTTTCTCTCAACAATTTCTTCTAGAGTTCTCTCTTTTAACATACATCGAGA	276
Db	291	TCCACTTTTACCTCTCATGACGATTTCTCCCAATGTTCTCCCTTCAACATATTGAAG	350
Qy	277	AAATCTACGGTGCATCAATTTCTGGTTGGTTCGTCGAATTTCCGGTTAACGGTAGCC	336
Db	351	AAATCTAATGTTCAATATTGTTTATATAGTTTGGACCAACGGCTGTGTACCACTTCT	410
Qy	337	GATCTGATTTGATCAGAGAG--TCTTCTTAAGTGTGATTTCTTACAGAGAGATGAA	393
Db	411	GATCCAGCTTTAATAGGATATATTCCTCTGAAATCAATATTTTACGAAAAATAG	470
Qy	394	GCTCAACCTTTGGTTAAACAATTGAAGGCATGACATACTTGTCTCAAAGTGAAAA	453
Db	471	TGCGCGGCGCTGTCAAGAGCTTGAAGTATGTTTGTGTGCTCCTCAAGGTAAAA	530
Qy	454	TGGGCTCATATCGAAAAATCATTTAGCCCTACITTTCAATATGAGAAATCTTAAGTTCTT	513
Db	531	TGGGCTCACCATAGGAAGTATTACTCTCATTTCTACATATAAATCTCAGGGCTAATG	590
Qy	514	GTAACAGTTGTGTTAAGATGTGACTATATGGTGAATAAATGSGTCCGATAGATATCA	573
Db	591	ATTCCCATATGGAGAGACATGAAGAGATGTTAACAATATGTCGCAAAATGTCTAAC	650
Qy	574	GAAACGCTGAAGTTAGGTTAGATGTCATGATGAGTGTTCAGATTTTGACTGAGATGTT	633
Db	651	GCCAGCGGAAAGTAGAGATGGAATGTGCAGAAATGTTCTCAACATTAGAGAGATGTC	710
Qy	634	ATTATGTAACACGCTTTTGAAGTATGCTATGAAGATGTTGAGACAGATTTTGCATCTCA	693
Db	711	ATTACGCGATATGTGTGGAAACGCTACGAAAGATGTAAGCCATCTTATGATTACA	770
Qy	694	GCTCAACAATGCTTTCTTGTCTGAAGCTTTGAAGAGCTTCAATCTCTGCTATAGA	753
Db	771	GCAACAACAATGTTTATGCCACCGAAGCTTATCAAAAAGATTCATCCCGGATACAG	830
Qy	754	TTTTTTCGACAGAGGGAATTGAAGTCTCGAAGTTAGCAAGAGATAGGAAGTGTG	813
Db	831	TTTTTGCTATAAAAAGATAGATATGTTGGAGATTGATTAACAAGTGTGTAATCT	890
Qy	814	TGTGTAAAGCTGATAGAGCGCGAG 839	
Db	891	TTATATGAGCTGATGAAAGAGAG 916	

RESULT 5
BU89.1384
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

614 bp mRNA linear EST 17-OCT-2002
BU89.1384
P049601 Populus petiolaris cDNA library Populus tremula cDNA 5 prime,
mRNA sequence.
BU89.1384
BU89.1384.1 GI:24102449
EST
Populus tremula
Populus tremula
Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; euroids; Malpighiales; Salicaceae; Salicaceae; Populus
1 (bases 1 to 614)
Umeberg, P., Bhalerao, R.R., Jansen, S. and Steyer, F.
The poplar tree transcriptome: Analysis of expressed sequence tags
from multiple libraries
Unpublished. (2002)
Contact: BHALERAO RUPALI R.

FEATURES

	source	1..614
	/organism="Populus tremula"	
	/mol_type="mRNA"	
	/db_xref="taxon:11636"	
	/tissue type="petioles"	
	/clone_id="Populus petioles cDNA library"	
ORIGIN		
Query Match	19.7%; Score 307.8; DB 13; Length 614;	
Best Local Similarity	72.4%; Pred. No. 3e-77;	
Matches 399; Conservative	0; Mismatches 152; Indels	Gaps 0;
QY	1013 TCTTGTCTTCATGCAACCCGGAGTGCGAGGCCAAGCAGTGATAGAGTCCGAGGGCT	1072
Db	1 TCCTACTGCATCATCTCCACAGTGGAGGCTCAGCAAGTAGAGAGGCTTTGAAGGCT	60
QY	1073 GCGGCTCACGATGATGCTCCTACCAAGAGCCATGCTTAAGCTTAAACGTTGAGATGA	1132
Db	61 GTGATCACTAGTAGACGCCACCAGGATATGTTCAAGGCTTAAGCGCTGACATGA	120
QY	1133 TCTTGAAGAGCTCTTTAAGTTGTATCCACAATAGTACTGACATTGAGACGCGCTAAT	1192
Db	121 TCTCTAAAGAATTCCTTCGCGCTGATCCCAACAATGCCCACAAATCAGGCCGTTCAAG	180
QY	1193 CGAGATGTAGAGTAGAGAGGTACAAAATCCCATAGTGGACGAGAGCTTCTAATCCCATCA	1252
Db	181 CTGATGTGAGCTGGGGGGCTCAAAATGCCAAGTGGAGCGGAAATCTGATACCGATCC	240
QY	1253 TAGCGGTCCATCATGACCAAGCCAATTTGGGGTATATGAGTGAACGAATTCATCCAGCTC	1312
Db	241 TGGCCCTTCATCATGATCTAATCAATAAGGGGCATATGCAAAATGAAATCAACCTGCTC	300
QY	1313 GGTTTGGCGATGAGATGCGCGTGTCTGCCAAACACCCGCTTGACTCAATACCGTTTGCC	1372
Db	301 GCCTTCCGGATGTGTGTGCCCGCGCTGCCAAGATCAAGTGGCGTTCAATCCGTTCCGGC	360
QY	1373 TCGAGTTGTGATCATGATCATTTGTGCAATTTTGTCTTACTTACAGGCCMAATGGAACATCG	1432
Db	361 TCGAGTCCGCACTCGCATCGGCCMAAATTTAACCATCTTGGCAAGCCAACTTAAGCTCG	420
QY	1433 CAGTATAGATCCAAAGCTTCACTTTCACTTGTGCTCTACTTATCAGATGACCATCTACCG	1492
Db	421 CTATATGTGCTGAAGATTTCTTTCAAGATGGGCCCTGTAGTATGAACATGACCAACAG	480
QY	1493 TCGTATGTTGCTTTATCTTCAACATGTGACCAATCACTTCCGAGATTTGACCATTC	1552
Db	481 TCGTCACTGCTACTTACCCCCAGATGTGGCAACCAATCACTTCCAAACATCTGTCCATTC	540
QY	1553 ATGAGAGATTGA 1563	
Db	541 ATGAAATATCA 551	

LOCUS	1676 bp	MRNA	linear	HTC 17-OCT-2002
AY109087	AY109087	PCO14494	1676 bp	
DEFINITION	Zea mays POC14494 mRNA sequence.			
ACCESSION	AY109087			
VERSION	AY109087.1	GI:21212491		
KEYWORDS	HTC.			
SOURCE	Zea mays			
ORGANISM	Zea mays			
	Elkaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD clade; Panicoideae; Andropogoneae; Zea.			
REFERENCE	1 (bases 1 to 1876)			
AUTHORS	Hathir,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsett,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.			
TITLE	maize Mapping Project/duPont Consensus Sequences for Design of Overgo Probes			
JOURNAL	Unpublished (2002)			
REFERENCE	2 (bases 1 to 1876)			
AUTHORS	Coe,E.H.			

TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSU, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES
Source location/Qualifiers
1. 1876
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="maizedb:638539"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/Dupont Cornsensus Library"

/note="This sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Query Match 18.4% Score 287.2; DB 11; Length 1876;
Best Local Similarity 51.4%; Pred. No. 5.1e-71;
Matches 751; Conservative 0; Mismatches 693; Indels 18; Gaps 3;

94 TGGTGAAGACCAAGAAATGTAAGACATTTCTTAAACAAGAAATTCAGGTCCTCT 153
136 TGGTGAAGACCAAGAAATGTAAGACATTTCTTAAACAAGAAATTCAGGTCCTCT 255
154 TATCATTTCTTATGGAATGTTAAAGAACTTTGTAAGTAAGTCTTAAAGTCTTCT 213
256 TACCGCTCTCTCAACGGGAGCGTGAAGAAACCTCGCTCAACCGAAGACCGGACG 315
214 CATCTATGCGC---TTCTCTCACAATATCTCTCTAGAGTCTCTCTTTTACCATCA 270
316 AAGCACTGCGCTGGAGATCAGACATCATCCAGCGCGTACCATGTTCCACAC 375
271 TGGAGAAAATCTACGCTGCTACATTTGCTGTTGTTGCTCCACTTCCGCTTACG 330
376 GCGGTGAAGAGAAATGGAGCAAAATTCATTCACTTGCTTGGCCGATTCGAGGGTATA 435
331 GTAGCGATTCCTGATTGATCAGAGATCTTCTTAAGTCTAGTCTAGGAAATCTTAA 390
436 ATTCAAGACCCGGAATTAATGAGAGATTTGTCTTAAAGTTGGCCACTTTGGGAAA 495
391 GAGCTCAACCTTTGTTTAAACAATTGAAGCGGATGACTTAACTGCTCAAGGTGA 450
496 CCACGTGTTAGCGGTTGCGCAAGTTACTAGCAATGGCTCGCTAATCAATGAGGCG 555
451 AATGGGCTCATCATCGAAAATCATAGCCCTACTTTTCATATGAGAAATCTTAAAGT 510
556 AATGGGCAAAACAGAGAAATCTCAACCTGCTATTCACATGAGAAAATTAACGG 615
511 CTGTACACAGTGTGTGTAAGTGAATGATATGATGATGATTAATGATGATGAT 570
616 ATGCTGCTGTTATTTGCTACCTGTTGTCGATATGATTAACATGAGAAATTCATG 675
571 TGAAGAAACGTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 630
676 TCTTCCAGAAACCTTCTGAATGAGCGCTGGCTGAGTTCGAAATCTTACTGAGAT 735
631 GTTATTAGTAAACAGCTTTTGAAGTGAATGATGATGATGATGATGATGATGATGAT 690
736 GTTATCTCAAGAACTCTTTTGTGTAAGTGAATGATGATGATGATGATGATGATGAT 795
691 CAAGCTCAAAATGCTTTTGTGTAAGTGAATGATGATGATGATGATGATGATGATGAT 750
796 CAAGGAGAGAGGCTGTAAGCTTTTAAATCTTTTCAACAAATTTATATCCAGGCTAC 855

751 AGATTTTCCGACAGAGGAAATTTGAAGTCTCGAAGTTAGACAGAGATTAAGAG 810
856 TGGCTTTCGCCACCAAAATTAACAAGATGAAGAAATTTGATCGGAGATTCGTTAA 915
811 TCGTTTGAAGCTGATGAGCGCGGAGACAAACCGTATGATGAGAGAGAGAG 870
916 ATTCTGATGAAATTAATACAGAAAGAGAGGCTTTTATGATGATGAGAGAACTAT 975
871 TGTAGAGAGCGCGGAGAGATTTTGGCATTAATGATGAGCAAGAAAGATG---- 925
976 GATACCTGTTAGGCTTATTTGTTGATCAACATAGGAAATCAATGGAATGAGAAA 1035
926 -----TGACGTTTCAGGACATTTGAGAGATGTAAGCTTTTTCGCGGAGAA 978
1036 CTCGAAATGATACCGAAGATTAATACAGAGATGCAAGTTTATTTACTTTGAGGTATG 1095
979 CAGCAACTTCTAATCTGCTGAGCTGAGACACACATCTGCTATCCATGACCCGAGTGG 1038
1096 GAGCAACATCAGTCTGCTTACATGACCTTAATTTCTGTAAGCATGACCTCTGATGG 1155
1039 CAGGCAAGACATGATGAGTCTCAGGATCTGCGGCTCAGTATGTCCTTACCAAG 1098
1156 CAAAGACGCGAAGAAAGATTTTAAACACTTTGG---AATGGGTACACGACTTT 1212
1099 GACCATGCTTTAAGCTTAAACGTTGATGATCTTGAACAGATCTTAAAGCTTGTAT 1158
1213 GATTAATGAAACCGCTGAAGATTTGTAACCATATTTTATGAGTCTTAAAGTTGATC 1272
1159 CCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1218
1273 CCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1332
1219 ATCCATGATGAGCAGGATCTTAAATCCATGATGATGATGATGATGATGATGATGAT 1278
1333 TATCTTCAAGCGTGAAGCTTCTACTGCTCATCATCTTCAATCAATGATGATGATGAT 1392
1279 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1338
1333 TGGGAAAGAGCGAAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1452
1339 GCCAAACCCCGTTGCTTCAATCCGTTTGGCTGAGATGATGATGATGATGATGATGATGAT 1398
1453 ACCAAGATGAGCGCGCTTCTTCCATTTGATGATGATGATGATGATGATGATGATGATGAT 1512
1399 AATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1458
1513 AACTTGGTGTGAGAGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
1459 CACTTGGCTTACTTATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 1518
1573 GAGCTCATATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1632
1519 GGTGACCAATCATCTTCCGGA 1540
1633 GGTGCTCAATCATGATGANA 1654

RESULT 7
CG027129/c 698 bp DNA linear GSS 19-AUG-2003
LOCUS CHG117/c CHGA Cleome hassleriana genomic clone CHG117, genomic
DEFINITION survey sequence.
ACCESSION CG027129
VERSION CG027129.1 GI:33899285
KEYWORDS GSS.
SOURCE Cleome hassleriana
ORGANISM Cleome hassleriana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Cleome.
REFERENCE 1 (bases 1 to 698)
Town,C.D., Van Aken,S., Utecherback,T. and Fraser,C.M.

QY 1085 ATGTCCTACCAAGACCAATGCTTAAAGCTTAAACGTTGATGATCTTTGAAGAGT 1144
 DB 551 ATCCCTCATTAACACCAAAATTTCCAGAGCTTAAGAGCTGAGATGATTAATTAAGCAAT 492
 QY 1145 CTTTAAAGTTGATTCACCAATAGATGATGATGACGCGCTAAATCGATGTGAAG 1204
 DB 491 CTGTCGCTATATCCACCAAGAGTGGCAGCAATCAACAGACCAAGTTGATGACAT 432
 QY 1205 TAGAGAGGTACAAATCCCATGTGAGCAAGAGCTTCAATCCCATATAGCGGTCCATC 1264
 DB 431 TAGAGGATTTCACTTACCTAGAGGAGTGAATCTCTTAATACCAATGATGCAATTCATC 372
 QY 1285 ATGACCAAGCCATTTGGGGTATATACCTGACCAATTCATCCAGCTCGGTTGGGGAGT 1324
 DB 371 ATGATCAAACTATGAGGAGCAAGAGCTTAACGAATTCATCAAGCAAGATTTGGTCTAG 312
 QY 1325 GAGTCCGCGTCTGCCAAACACCCCGTTGCTTCAATACGCTTTGGCTCGAGATTGCTA 1384
 DB 311 GAGTGGCAACAGCAGCAAAACACCCCATGCGCTTTCATGCTTTGGCTCGAGCGCCGAC 252
 QY 1385 CATGATTTGTCAGATCTTGGCTATCTTCAGGCCAATTTGACACTGCTGTATGATTC 1444
 DB 251 GATGCGTGGCCAAATTTAGCAGTGTACAGCTTAATTAAGCAATGCTATGATCTTGC 132
 QY 1445 AAGGCTTACCTTCACTTGGCTCTCTATCTATACAGATGACCTACCGCTTATGTTGC 1504
 DB 191 AAGGCTTTTCGTTTATCTTTCTCCATTAACAGATGCTCCCATATTTGATGCTCT 132
 QY 1505 TTTATCTTCACATGCTGACCAATCACCTTCCGAGATTG 1545
 DB 131 TATGTCCAAATATGCTGCTCTATTAATTCACAAAGTTG 91

RESULT 9
 BH597310 817 bp DNA linear GSS 15-DEC-2001
 LOCUS BH597310
 DEFINITION BH597310 BOLT Brassica oleracea genomic clone BH597310, genomic survey sequence.
 ACCESSION BH597310 GI:17849762
 VERSION BH597310
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 817)
 AUTHORS Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished (2001)
 COMMENT Other GSSs: BH597310
 CONTACT: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seg primer: TR
 Class: sheared ends.

FEATURES

Source

1. 817
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BH597310"
 /clone_1="BH597310"
 /note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

ORIGIN

Query Match 16.6%; Score 259.6; DB 28; Length 817;
 Best Local Similarity 83.9%; Pred. No. 3e-63;
 Matches 317; Conservative 0; Mismatches 59; Indels 2; Gaps 2;
 QY 1122 GTTAGATATGATCTTGAAGAGTCTTTAAGTTGATTCACCAATATGATGATGATG 1181
 DB 441 GTTGGGATGATCTTGAAGAGTCTTTAAGTTGATTCACCAATATGATGATGATG 500
 QY 1182 ACGGCTAAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1241
 DB 501 ACGGCTAAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 560
 QY 1242 AATCCCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1301
 DB 561 GATCCCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 620
 QY 1302 CATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1361
 DB 621 CATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 680
 QY 1362 ACCGTTGGCTCGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1421
 DB 681 ACCGTTGGCTCGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 739
 QY 1422 ATTGACTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1480
 DB 740 ACTGACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 799
 QY 1481 ATGACCTACCGTCTCTTA 1498
 DB 800 ACGACCTACCGTCTCTTA 817

RESULT 10
 BQ297122 493 bp mRNA linear EST 16-MAY-2002
 LOCUS BQ297122
 DEFINITION BQ297122 493 bp mRNA linear EST 16-MAY-2002
 ACCESSION BQ297122
 VERSION BQ297122.1 GI:20812644
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 REFERENCE 1 (bases 1 to 493)
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, J., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R., Riltter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished (1999)
 COMMENT Public Soybean EST Project
 CONTACT: Shoemaker R/Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available through: Reggen, Invitrogen Corp. 2130
 South Memorial Parkway Huntville, AL 35801 For further information
 call: (800)-533-4363 or contact: cst@reggen.com web site:
 www.reggen.com
 Seg primer: -40RP from Gibco
 High quality sequence stop: 425.
 Location/Qualifiers
 1. 493
 /organism="Glycine max"

FEATURES

Source

/mol_type="mRNA"
/db_xref="taxon:3947"
/clone="SOBESAN CLONE ID: Gm-cl054-6364"
/issue_type="leaf", 3 week old, greenhouse grown"
/lab_host="DH10B"
/clone_lib="Gm-cl054"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI. The Harosoy NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from leaf tissue at various developmental stages of 3 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(AT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."

ORIGIN

Query Match 16.4%; Score 256; DB 13; Length 493;
Best Local Similarity 70.3%; Pred. No. 2,5e-62;
Matches 343; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 1057 GAGGTCTCAGGCTCGCGCTCAGCGTGAATGCCACAGAGCATGTGCTTAAGCTT 1116
Db 2 GAGGTCTCAGGCTCGCGCTCAGCGTGAATGCCACAGAGCATGTGCTTAAGCTT 61
QY 1117 AAAACCTTGAGTGAATCTTGAACGAGTCTTTAAGTTGATCCACCAATAGTACTACG 1176
Db 62 AGGACGCTGAGCATGATTTGAAAGATCAGTCACTGACCTCCGACGATCCGACCA 121
QY 1177 ATTGACGCGCTAATATGATGTAAGAGGCTAGCAAAATCCCATGTCGCGAG 1236
Db 122 ATCAGAGGCGCAAGGAGGAGCTGATTTGGGGGATACAAATACACGTGGAGCGAG 181
QY 1237 CTTCTAATCCCAATCATAGCGGTCCATCATGACCAACCAATTTGGGTAATGACGTGAAC 1236
Db 182 CTGTGATTCACATCCGTGCGCGTTCATCAGATCAAGCAATATGGGGAGAACATGTGAAC 241
QY 1297 GAATTCATCCAGTCTGCTTGGCGATGAGTCCCGGCTGCTCCCAAAACCCCGTTGGC 1356
Db 242 GAGTTCAATCCGGGGCGTTTCAGCGATGATGTCGCGCGGGGAGGATCCATTTGGCG 301
QY 1357 TTCTAATCGTTTGGCTCGGAGTTGCTAATGATGATGATGATGATGATGATGATGAT 1416
Db 302 TTCTAATCGTTTGGCTCGGAGTTGCTAATGATGATGATGATGATGATGATGAT 361
QY 1417 GCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1476
Db 362 ACAAACCTGCGCGCTGAGATCACTGACAGCGTTTAACTTAACTTAACTTAACTTAACT 421
QY 1477 CAGATGACCTACCGTCTTATGTTGTTTATCTTATCTTATCTTATCTTATCTTATCTTAT 1536
Db 422 CAGATGACCTACCGTCTTATGTTGTTTATCTTATCTTATCTTATCTTATCTTATCTTAT 481
QY 1537 CGGAGATT 1544
Db 482 CAACATT 489

RESULT 11
B62694/c 303 bp DNA linear GSS 21-NOV-1997
LOCUS B62694
DEFINITION T17P24TF.1 TAWU Arabidopsis thaliana genomic clone T17P24, genomic
survey sequence.
ACCESSION B62694
VERSION B62694.1 GI:2629456
KEYWORDS GSS.

SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustoids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
AUTHORS Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter,J.C.
TITLE A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genome Sequencing. Update 3
JOURNAL Unpublished (1997)
COMMENT Other_GSS: T17P24TF T17P24TRB
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@igr.org
Seq primer: M13-21
Class: BAC ends
High quality sequence stop: 303.
FEATURES
source
1..303
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T17P24"
/sex="hermaphrodite"
/clone_lib="TAWU"
/note="Vector: pBluescript II SK+; Site 1: HindIII; Site 2: HindIII; Produced by Rod Wing"

ORIGIN

Query Match 15.6%; Score 244; DB 28; Length 303;
Best Local Similarity 98.0%; Pred. No. 5,5e-59;
Matches 247; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1312 CGGTTTGGCGATGAGTGGCGGCTGCTGCGCAACACCCCGTTGGCTTATACCGTTGGC 1371
Db 303 CGTTTCCGATGAGTGGCGGCTGCTGCGCAACACCCCGTTGGCTTATACCGTTGGC 244
QY 1372 CTGCGAGTTCATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1431
Db 243 CTGCGAGTTCATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 184
QY 1432 GCTGTATGATTCACAGCTTTCACCTTTCACCTTTCACCTTTCACCTTTCACCTTTCAC 1491
Db 183 GCTGTATGATTCACAGCTTTCACCTTTCACCTTTCACCTTTCACCTTTCACCTTTCAC 124
QY 1492 GTCTTATGTTGCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTAT 1551
Db 123 GTCTTATGTTGCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTAT 64
QY 1552 CATGAGATTGA 1563
Db 63 CATGAGATTGA 52

RESULT 12
CG112869/c 834 bp DNA linear GSS 20-AUG-2003
LOCUS CG112869
DEFINITION PUIA51TB ZM_0.6.1.0_KB Zea mays genomic clone ZM05B053805,
genomic survey sequence.
ACCESSION CG112869
VERSION CG112869.1 GI:33996306
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

Clade: Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 834)
 White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
 Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
 Bennetzen, J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other GSSs: PUIK51TD
 Contact: Cathy White, cwhite@maizegenetics.org
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: white@maizegenetics.org
 Seq primer: TR
 Class: sheared ends.

FEATURES
 source
 1. 834
 /organism="Zea mays"
 /mol_type="Genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMB180538J05"
 /clone_id="ZM_0.6_1.0_KB"
 /note="Vector: pCR4-TOP0; Site_1: EcoRI; 0.6-1.0 kb high
 Cot selected genomic DNA library"

ORIGIN

Query Match 15.5%; Score 242.4; DB 29; Length 834;
 Best Local Similarity 61.9%; Freq. No. 2.9e-58;
 Matches 384; Conservative 0; Mismatches 236; Indels 0; Gaps 0;

926 TGACGGTTCAGACGCTTGTGAGAGAGGTAAAGCTTTTCTTCCGCCGAGAAACAGACAA 985
 631 TGGCGCGGAGAGATCATCGAGAGATCAAGAACTTCTTCCGCCGAGAGAGAGCG 572
 986 CTTTAAATCTGTGACGCTGAGACGACCACTTCTTATCATGACACCGGAGTGGCAGGCCA 1045
 571 TCACCAAGTTTGTCTCCTCGGCGCACCGCTGCGCTTACGATGACAGAAATGGCAAGACC 512
 1046 AAGCAGGTGATGAGTCTCTCAGGCTCTGCGCTCAGCGATGATGCTTACCAAGACCATG 1105
 511 GCGCTCGGAGAGATCATGAGATCTGCGGACACCGGCGCTCCACCAAGAGACACA 452
 1106 TCGTTAAGCTTAAACGTTGAGTATGATCTTGAACGAGTCTTTAAGTTGATTCACCAA 1165
 451 TACCAAGGCTCAAGAGCGCTCGGAGTGTCAACAGACGCTCAGGCTTACCCGCGG 392
 1166 TGTAGCTACGATTCGAGCGGCTTAATGAGATGAGAGCTAGAGGCTACAAATCCCAT 1225
 391 CGGTGCGCATGATCCGAAAGCGGAGCGGAGCTGAGCTGCGGCTGCTGTCGCG 332
 1226 GTGGACGAGAGCTTAAATCCCATATAGCGGCTCATGATGACCAAGCATTTGGGCTA 1285
 331 CGGGACGAGATCATGATCCCATATGCGGCTACACAGACGCGGCTGTGGGCG 272
 1286 ATGAGCTGACGATTCATTCAGCTCGGTTTGGGAGTGGAGTCCGCGGTCTGCCAAAC 1345
 271 ACGAGCGCACGAGATTCAGCCCGGCGGCTTGGCGAGATGATGACCGCGCGCGGCG 212
 1346 ACCCGGTTGCTTCAATACCGTTTGGCTTCGAGATTCGATGATGATGATGATGATGAT 1405
 211 ACGAGATGCGCTTCTTCCGTTCCGCGGCGGCGCGGCTGATGATGATGATGATGATGAT 152
 1406 CTAATCTTACGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1465
 151 CGCTCATGAGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 92
 1466 CTCCTACTTATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1525
 91 CGCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 32

Db 31 CGGTACATCTCCGCCCATG 12

RESULT 13
 AI104671
 LOCUS
 DEFINITION
 Zea mays PC0144495 mRNA sequence.
 AY104671
 ACCESSION
 AY104671.1 GI:21207749
 VERSION
 KEYWORDS
 mRNA
 SOURCE
 Zea mays
 ORGANISM
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade: Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 1567)

REFERENCE
 AUTHORS
 Hainey, C.F., Doan, M., Mao, G.H., Vogel, J.M., Whitsitt, M.S.,
 Arthur, L.W., Hanaley, M., Morgante, M. and Tingey, S.V.
 Maize Mapping Project/Dupont Consensus Sequences for Design of
 Overgo Probes
 Unpublished (2002)
 2 (bases 1 to 1567)
 Coe, E.H.
 Direct Submision
 Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA

COMMENT

If you are interested in getting corresponding physical clones,
 these are publicly available from ZMDB and may be found by BLAST
 searching at MSU, maizegap.org; ZMDB, www.zmdb.iastate.edu; TIGR,
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat
 Schnable, Iowa State, then clones may be requested from ZMDB:
 www.zmdb.iastate.edu.

FEATURES
 source
 1. 1567
 /location/Qualifiers
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="maizedb:638540"
 /db_xref="taxon:4577"
 /clone_id="Maize Mapping Project/Dupont Consensus
 Library"

/note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont contigs; this resource was
 assembled by Dupont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

ORIGIN

Query Match 15.0%; Score 234; DB 11; Length 1567;
 Best Local Similarity 50.2%; Freq. No. 1.1e-55;
 Matches 632; Conservative 0; Mismatches 620; Indels 6; Gaps 2;

284 ACGGTGTACATTTCTGTGTTTGGTTCGATCCACTTCCGTTAAACGTTAGCGGATCTG 343
 101 ACGGAGAGACGAGATGCTCGGTTGCGGCCACGCCAAGAGTACATGTCGATCCCG 160
 344 ATTGATCAGAGAGATCTTCTTAAGCTGATGATCTCAGAGAGATGAGATGATGATGATGAT 403
 161 ACGTGGCAGAGAGATGCTGCTCAAGATGCGGCTTCAAGAGATGATGATGATGATGATGAT 220
 404 TGGTTAAACAACTTGAAGGCGATGATCTTAACTTCAAGATGATGATGATGATGATGATGAT 463
 221 CGCTGTCCAGAGATGCTGCGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 280
 464 ATCGAAATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 523
 281 ACGGAGAGATGCTCAACCTTGTGCTTCAATCTTGAAGATGATGATGATGATGATGATGAT 340
 524 TGTGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 580
 341 TTCTACGCTGCTGAGAGAGCTGTGAGCAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 400

Mon May 24 08:10:09 2004

us-09-992-901-1.rst

Page 11

KEYWORDS
SOURCE

Zea mays
Zea mays

ORGANISM

REFERENCE

White, J. A., Quackenbush, J., Van Aken, S., Utecht, T.,
Bennick, A., Fraser, C. M., Budiman, M. A., Bedell, J. A., Rohlfing, T.,
Clerk, R. W., Numbers, A., Robbins, D. and Lakey, N.

AUTHORS

TITLE
JOURNAL

Unpublished (2002)
Contact: Cathy Whiteleaw

COMMENT

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whiteleaw@tigr.org

Seq primer: TF

Class: sheared ends.

FEATURES

source

1..815

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBMA0392E01"

/clone_lib="ZM 0.7-1.5 KB"

/note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match

14.0%; Score 219.6; DB 28; Length 815;

Best Local Similarity

57.3%; Pred. No. 1.1e-51;

Matches 442; Conservative

0; Mismatches 319; Indels 11; Gaps 2;

727 CAAAAGCTTCATTCCTCGGTATAGATTTTCCGACAGAGGGAATTGAAGTCTCG 785
773 CAGCAAGCTTCCTCCCGGCTACAGGTTCTCTCCACAGAGGAACAAGCGCTGTG 714
786 GAAGTTAGACAGAGAGATAGAGAGTGTGTTGAAGTATAGAGCGCGAGACAAAA 845
713 GCAAGTTAGACAGAGAGATAGAGAGTGTGTTGAAGTATAGAGCGCGAGACGCG 654
846 CGCTATAGATGAGAGAGAGAGAGATAGAGAGCGCGCGAGAGATTGTTGGATT 905
653 CGGACAGCGGTATCCGACACAGAGCGCGGAGACAGAGCGCGCGCGGATGA 594
906 AATGATGA-----GGCAAGATGACGTTCAAGACATTTGAGAGAGTGA 955
593 GGAATTCATGAGCTTCATGAGCGCGCGGATGAGCGCGAGATCATGAGGAATCCA 534
956 AAAGCTTTTCTTCGCGCGGAGAAACAGACAACTTTAATCTGCTGAGCGATCATCT 1015
533 AGAATCTTCTTCGCGCGGAGAGAGAGCTCACCAAGTTGCTCAGCTGAGCGCA 474
1016 TGCTATCATGACCGGAGAGTGGAGCGCGCAAGACAGTATGAGTCTTCAGGATCG 1075
473 CGCTAGCCATGACAGAGAGATGAGAGCGCGCGCTCGGAGAGATCATGAGAGTCT 414
1076 GCTCAGTATGCTCCCTACAGAGACATGCTGTTAAGCTTAAACGTTAGATGATCT 1135
413 GAGCGCGGCGCGTCCCAACAGAGACATACCAAGCTCAAGACGCTCGGAGTGTCA 354
1136 TGAAGAGCTTTAAGGTTGATTCACCAATAGTAGCTAGATTCGACGCGCTAAATCG 1195
353 TCAGAGAGAGCTTCAGGCTTACCGCGCGGAGTGCATATCCGAGAGCGAGCGGG 294
1196 ATGTGAAGCTAGAGAGGTACAAATCCCATGTGAGCGAGAGCTTCTAATCCCAATCA 1255
293 AGGTGAAGCTCGCGCGGCTGCTGAGTGGCGGAGAGATCATGATCCCATCATAG 234

233 CGGTACACACAGACCGCGCGTGTGGGCGAGACGCCACGAGTTGACCCCGCGCT 174
1316 TTGGAGATGAGAGTCCGCGTGTGCGCAACACCCGTTGCTTACCGTTGAGCTCG 1375
173 TCGGAGATGAGTGTGACCCCGCGCGGAGAGATGCGCTTCCGTTGCGCGCG 114
1376 GAGTGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1435
113 GCGCGCGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 54
1436 TAATGATCAAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1487
53 CGGTGTGACAGCGCTGCAAGTTCCGCTTCCGCGGAGTACGATGACGCGCC 2

Search completed: May 22, 2004, 20:13:11
Job time : 4347 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2004, 15:23:56 ; Search time 133 Seconds

(without alignments)
6521.716 Million cell updates/sec

Title: US-09-992-901-1

Perfect score: 1563
Sequence: 1 atggaggaagaagtagcag.....tgaccaatcatgagatga 1563

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6C.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6D.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/6E.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	1563	100.0	1563	4	US-09-527-073-1
2	270.4	17.3	1584	4	US-09-351-229-1
3	184	11.8	588	4	US-09-527-073-16
4	75.8	4.8	612	4	US-09-615-192A-101
5	75.8	4.8	612	4	US-09-169-789-101
6	73.4	4.7	398	4	US-09-615-192A-169
7	73.4	4.7	398	4	US-09-169-789-169
8	68.6	4.4	285	4	US-09-313-294A-2775
9	53	3.4	7218	1	US-08-232-463-14
10	48	3.1	933	4	US-09-118-554-55
11	48	3.1	933	4	US-09-118-627-55
12	48	3.1	933	4	US-09-602-877A-55
13	48	3.1	940	4	US-09-602-877A-102
14	46.2	3.0	1708	3	US-08-991-677-1
15	45.8	2.9	4403765	3	US-09-103-840A-2
16	45.8	2.9	4411529	3	US-09-103-840A-1
17	44	2.8	2382	4	US-09-023-655-1436
18	43.6	2.8	1512	4	US-08-277-031B-4
19	43.6	2.8	2059	4	US-09-023-655-1062
20	43.6	2.8	2759	4	US-09-144-367-1
21	43	2.8	1515	4	US-09-583-447A-3
22	41.6	2.7	1725	4	US-08-882-164D-31
23	41.6	2.7	4411529	3	US-09-103-840A-1
24	41.4	2.6	1893	1	US-08-533-065B-1
25	41.4	2.6	2327	4	US-09-852-067-1
26	40.2	2.6	1634	4	US-09-126-420A-2
27	40.2	2.6	1737	4	US-09-126-420A-1

28	40	2.6	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
29	38.4	2.5	1932	4	US-09-518-386B-2	Sequence 2, Appl1
30	37.8	2.4	936	2	US-08-557-309B-9	Sequence 9, Appl1
31	37.8	2.4	936	3	US-08-834-306-9	Sequence 9, Appl1
32	37.8	2.4	936	3	US-08-993-674A-9	Sequence 9, Appl1
33	37.8	2.4	936	3	US-09-256-976-9	Sequence 9, Appl1
34	37.6	2.4	289	3	US-08-007-605-17	Sequence 17, Appl1
35	37.6	2.4	289	3	US-09-244-796-17	Sequence 17, Appl1
36	37.6	2.4	7218	1	US-08-232-463-14	Sequence 14, Appl1
37	37.4	2.4	1665	3	US-08-881-784-8	Sequence 8, Appl1
38	37.4	2.4	1665	3	US-09-292-768-3	Sequence 3, Appl1
39	37.4	2.4	1665	3	US-09-292-768-67	Sequence 67, Appl1
40	37.4	2.4	1665	3	US-08-292-768-69	Sequence 69, Appl1
41	37.4	2.4	3969	4	US-08-518-386B-4	Sequence 4, Appl1
42	37.2	2.4	274	4	US-09-313-294A-1017	Sequence 1017, Ap
43	37.2	2.4	1707	4	US-09-023-655-1060	Sequence 1060, Ap
44	37	2.4	1192	4	US-09-583-447A-9	Sequence 9, Appl1
45	37	2.4	1633	4	US-09-583-447A-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-09-527-073-1
Sequence 1, Application US/09527073
Patent No. 6534313
GENERAL INFORMATION:
APPLICANT: Michael M. Neff
APPLICANT: Joanne Chory
TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING
FILE REFERENCE: SALKINS 024A
CURRENT APPLICATION NUMBER: US/09/527, 073
CURRENT FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: US 60/124570
PRIOR FILING DATE: 1999-03-16
PRIOR APPLICATION NUMBER: US 60/170,931
PRIOR FILING DATE: 1999-12-14
PRIOR APPLICATION NUMBER: US 60/172,832
PRIOR FILING DATE: 1999-12-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1563
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-09-527-073-1

Query Match 100.0%; Score 1563; DB 4; Length 1563;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAGGAGGAAAGTATGAGTGTCTCTGTTATGAGGAGCAAGAAAGATGAGAA	60
DB	1	ATGAGGAGGAAAGTATGAGTGTCTCTGTTATGAGGAGCAAGAAAGATGAGAA	60
QY	61	AGCTCTGTAAGTGAAGGAGTATCTCTGTTATGAGGAGCAAGAAAGATGAGAA	120
DB	61	AGCTCTGTAAGTGAAGGAGTATCTCTGTTATGAGGAGCAAGAAAGATGAGAA	120
QY	121	CATTTCTTAACCAAGGATTCGAGTCTCTCTTATATTTCTTCATCGAAATGTAA	180
DB	121	CATTTCTTAACCAAGGATTCGAGTCTCTCTTATATTTCTTCATCGAAATGTAA	180
QY	181	GAACTGTGGAAGATGCTTAAGGCTTCTTCATCCATGCTTCTCTCAATATT	240
DB	181	GAACTGTGGAAGATGCTTAAGGCTTCTTCATCCATGCTTCTCTCAATATT	240
QY	241	CTTCTAGAGTCTCTCTTTTACCATCTGAGAAATCTACGGTCTACATTTCTG	300
DB	241	CTTCTAGAGTCTCTCTTTTACCATCTGAGAAATCTACGGTCTACATTTCTG	300

Db 241 CTTCCTAGAGTTCTCTCTTTTACATCATGAGAGAAAATCTACGGTGTACATTTCTG 300
Qy 301 GTTGTGTTGGTCCAACTTCCGGTTAAACGGTAGCCGATCTGTATTTGATCAGAGATC 360
Db 301 GTTGTGTTGGTCCAACTTCCGGTTAAACGGTAGCCGATCTGTATTTGATCAGAGATC 360
Qy 361 TTCTTAAGTCTGAGTTCTACAGAGAAATGAAGTCAACCTTTGGTTAAACAATTGAA 420
Db 361 TTCTTAAGTCTGAGTTCTACAGAGAAATGAAGTCAACCTTTGGTTAAACAATTGAA 420
Qy 421 GGGGATGGACTACTAGTCTCAAGGTGAAAAATGGGCTCATCATCGAAAAATCATTAAC 480
Db 421 GGGGATGGACTACTAGTCTCAAGGTGAAAAATGGGCTCATCATCGAAAAATCATTAAC 480
Qy 481 CCTACTTTTCATATGAGAAATCTTAAGTGTGTTACAGAGTTGTGTAAGAGTGTACT 540
Db 481 CCTACTTTTCATATGAGAAATCTTAAGTGTGTTACAGAGTTGTGTAAGAGTGTACT 540
Qy 541 GATATGATGATTAATGATCCGATTAAGTTATCAGAAAACGGTGAAGTTAGGTATGATGC 600
Db 541 GATATGATGATTAATGATCCGATTAAGTTATCAGAAAACGGTGAAGTTAGGTATGATGC 600
Qy 601 TATGATGTTTCAAGTTTGAAGTGAAGTGTATTAAGTGAAGAGTGTGGAAGTAC 660
Db 601 TATGATGTTTCAAGTTTGAAGTGAAGTGTATTAAGTGAAGAGTGTGGAAGTAC 660
Qy 661 TATGAGATGTCGAGACAGTTTTCAGTTCAAGTCAACAATGCTTTTGTGTCTGAA 720
Db 661 TATGAGATGTCGAGACAGTTTTCAGTTCAAGTCAACAATGCTTTTGTGTCTGAA 720
Qy 721 GCTTTTCAAAAAGTCTTCATCTCTGCTATTAAGTTTTCGACAAGAGGAAATTGAG 780
Db 721 GCTTTTCAAAAAGTCTTCATCTCTGCTATTAAGTTTTCGACAAGAGGAAATTGAG 780
Qy 781 TCTCGGAGTTTACAGAGAGATTAAGAGTGTGTTGAAGTGTATGAGAGTGTGAG 840
Db 781 TCTCGGAGTTTACAGAGAGATTAAGAGTGTGTTGAAGTGTATGAGAGTGTGAG 840
Qy 841 CAAAACGCTATTAAGTGAAGAGGAGAAATGTAAGAGCCGGCCGCAAGGATTTGTTG 900
Db 841 CAAAACGCTATTAAGTGAAGAGGAGAAATGTAAGAGCCGGCCGCAAGGATTTGTTG 900
Qy 901 GGAATTAATGATTAAGAGGAAATGTAAGAGTGTGAGATTTGTAAGAGTGTGAAAGC 960
Db 901 GGAATTAATGATTAAGAGGAAATGTAAGAGTGTGAGATTTGTAAGAGTGTGAAAGC 960
Qy 961 TTTTCTTCCCGCGGAAACAGACAACTTCTAATCTGCTGACGTGAGACATCTTGTCTA 1020
Db 961 TTTTCTTCCCGCGGAAACAGACAACTTCTAATCTGCTGACGTGAGACATCTTGTCTA 1020
Qy 1021 TCCATGACACCGGAGTGGCAGGCAAGACGATGAGAGTCTCTAGAGTCTGCGGCTCA 1080
Db 1021 TCCATGACACCGGAGTGGCAGGCAAGACGATGAGAGTCTCTAGAGTCTGCGGCTCA 1080
Qy 1081 CGTATGTCCTTACCAAGAGCATGTCGTTAACTTAAACGTTAGTATGATCTTGAAC 1140
Db 1081 CGTATGTCCTTACCAAGAGCATGTCGTTAACTTAAACGTTAGTATGATCTTGAAC 1140
Qy 1141 GAGTCTTTAAAGTTGTTTCCACCAATGATGATGATGATGATGATGATGATGATGATG 1200
Db 1141 GAGTCTTTAAAGTTGTTTCCACCAATGATGATGATGATGATGATGATGATGATGATG 1200
Qy 1201 AAGCTAGAGAGGTTCAAAATCCATGTCGACGAGACTTCAATCCCAATCAATAGAGTC 1260
Db 1201 AAGCTAGAGAGGTTCAAAATCCATGTCGACGAGACTTCAATCCCAATCAATAGAGTC 1260
Qy 1261 CATCATGACCAAGCATTTTGGGGTATGAGAGTGAAGATTCATTCACACTGGTTTGGC 1320
Db 1261 CATCATGACCAAGCATTTTGGGGTATGAGAGTGAAGATTCATTCACACTGGTTTGGC 1320
Qy 1321 GATGAGAGTGGCGGTGTCGCAAAACACCCGTTGCTTCAACGTTTGGCTCGAGTT 1380
Db 1321 GATGAGAGTGGCGGTGTCGCAAAACACCCGTTGCTTCAACGTTTGGCTCGAGTT 1380

Qy 1381 CGTACATGATTTGTCAGAAATCTTGTATTAATCTTACGCGCAAAATGACATCGCTATATG 1440
Db 1381 CGTACATGATTTGTCAGAAATCTTGTATTAATCTTACGCGCAAAATGACATCGCTATATG 1440
Qy 1441 ATCCAGCGCTTACCTTCTTCACTTGGCTCTTACTTATGAGATGACCTTACCGCTTATG 1500
Db 1441 ATCCAGCGCTTACCTTCTTCACTTGGCTCTTACTTATGAGATGACCTTACCGCTTATG 1500
Qy 1501 TTGCTTTATCTTCAACATGATGACCAATGACCTTCCGAGATTAACCAATCATAGAGAT 1560
Db 1501 TTGCTTTATCTTCAACATGATGACCAATGACCTTCCGAGATTAACCAATCATAGAGAT 1560
Qy 1561 TGA 1563
Db 1561 TGA 1563

RESULT 2
US-09-351-229-1
; Sequence 1, Application US/09351229
; Patent No. 6380465
; GENERAL INFORMATION:
; APPLICANT: Barrett, Michael
; TITLE OF INVENTION: Cytochrome P450 Enzymes and Related Compounds and
; TITLE OF INVENTION: Methods
; FILE REFERENCE: P-1049
; CURRENT APPLICATION NUMBER: US/09/351, 229
; EARLIER FILING DATE: 1999-07-11
; EARLIER APPLICATION NUMBER: 60/092, 596
; EARLIER FILING DATE: 1998-07-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1584
; TYPE: DNA
; ORGANISM: Zee mays
US-09-351-229-1

Query Match 17.3%; Score 270.4; DB 4; Length 1584;
Best Local Similarity 51.5%; Pred. No. 1.1e-74; Indels 15; Gaps 2;
Matches 680; Conservative 0; Mismatches 626;

Db 228 CTCTCAATATTTCTTCTAGAGTCTCTCTTTTACATCACTGAGAGAAATCTACGG 287
Qy 228 CTCTCAATATTTCTTCTAGAGTCTCTCTTTTACATCACTGAGAGAAATCTACGG 287
Db 249 CTCCACGACATCACCCACCGGTGACCCATGATCAGACACATCAAGAAATCGG 308
Qy 249 CTCCACGACATCACCCACCGGTGACCCATGATCAGACACATCAAGAAATCGG 308
Db 288 TGCTACATTTCTGTTGGTTGGTTCGCACTTCCGGTTAACGGTAGCCGATCCTGATTT 347
Qy 288 TGCTACATTTCTGTTGGTTGGTTCGCACTTCCGGTTAACGGTAGCCGATCCTGATTT 347
Db 309 GAAACTATGCTTCACTGCTTGGGCAACCAAGGATGATGATCTGACCAAGATT 368
Qy 309 GAAACTATGCTTCACTGCTTGGGCAACCAAGGATGATGATCTGACCAAGATT 368
Db 348 GATCAGAGAGATCTTCTTAAGTCTGAGTTCTACAGAGATGAGTCAACCTTTGTT 407
Qy 348 GATCAGAGAGATCTTCTTAAGTCTGAGTTCTACAGAGATGAGTCAACCTTTGTT 407
Db 369 AGTCAGAGAGTCTCTTAAGTCTGAGTTCTACAGAGATGAGTCAACCTTTGTT 428
Qy 369 AGTCAGAGAGTCTCTTAAGTCTGAGTTCTACAGAGATGAGTCAACCTTTGTT 428
Db 408 TAAACACTTGAAGGAGTGAATCTTAACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 467
Qy 408 TAAACACTTGAAGGAGTGAATCTTAACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 467
Db 429 TGGAGAGTGTAGCCCAAGCGGCTTGTAAATGATGATGATGATGATGATGATGATGATG 488
Qy 429 TGGAGAGTGTAGCCCAAGCGGCTTGTAAATGATGATGATGATGATGATGATGATGATG 488
Db 468 AAAAATCATTAAGCCCTTCTTCAATGATGAGATCTTAAGTGTCTTGAACAGTTGTT 527
Qy 468 AAAAATCATTAAGCCCTTCTTCAATGATGAGATCTTAAGTGTCTTGAACAGTTGTT 527
Db 489 GAGATTTCTTAATCTGATTTCAACAGGAAATTAAGGAGATGATGATGATGATGATGATG 548
Qy 489 GAGATTTCTTAATCTGATTTCAACAGGAAATTAAGGAGATGATGATGATGATGATGATG 548
Db 549 TACCTGCTGATTAATGATTAATTAATGATGATGATTAATTAATGATGATGATGATGATG 608
Qy 549 TACCTGCTGATTAATGATTAATTAATGATGATGATTAATTAATGATGATGATGATGATG 608
Db 588 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 647
Qy 588 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 647
Db 609 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668
Qy 609 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668
Db 648 TTTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 707
Qy 648 TTTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 707

Db 669 GTTTGGGACCACTATCAAGAAAGGAGAGATTTTGTAGCTACAGGAGAACTAGCTGA 728
 QY 708 TCTTTGTGCTGAAGCTTTTCAAAAAGCTTCAATTCCTGGCTATGATTTTTCGGACAG 767
 Db 729 AGCCTCAATCCAACTGTGTGACACATTTTATCCAGGCTATGTGTTCTTGGCCACAA 788
 QY 769 AGGGAATTTGAAGTCCGAAGTGTAGACAGAGATAGAAAGTCTTGTGTAAGCTAT 827
 Db 789 AAACAACAGAGAGAGAGACATGATGTAGAGATCCGCAAAATTCCTCGTGAATAT 848
 QY 828 AAGAGCGGCGAGACAAACGCTATATAGATGAGAGAGAGAGAGATGTAAAGG----- 879
 Db 849 TGGAGAGAGAGAGAGATGATTAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 908
 QY 880 -----CGGCGGCGAGAGATTTGTGGAGTTATGATGAGCAAAAGATGTAGCGTTCA 935
 Db 909 ATTACTGAGAGTCAACACAGAGCAATCAATGAGAAATCAAGCTCGGATTTGACACAGA 968
 QY 936 GGCATTTGTGAGAGAGTGAAGGCTTTTCTTCGCGGAGAAACAGCACTTCTAATCT 995
 Db 969 AGATGTGATGAGAGATGAGATTTTACTTTTCAAGTATGAGAGACATCAGTCTCT 1028
 QY 996 GCTGACGTGAGACCATTTGCTATCCATGACCCGAGAGTGGCAGGCGCAAGACGTGA 1055
 Db 1029 GCTTACTTGGACACTTATGTGTCTATGACATGACCCAGAGAGAGAGAGAGAGAGAG 1088
 QY 1056 TGAAGTCTCAGAGGTGTGGGCTCAGTATGCTCCCTACCAAGACATGCTCTTAAAGCT 1115
 Db 1089 AGAGGTTTGAAGCACTTTGG--AAGAACACACCAATATGATGATGATGAGAGAGAG 1145
 QY 1116 TAAACGTTGAGATGATCTTGAACGAGCTTTTAAAGTTGATTCACCAATATGATGAT 1175
 Db 1146 CAAGACTGATACCATGATCTTACATGAGGCTCTTGTGATGATGATGATGATGATGAT 1205
 QY 1176 GATTCAGACGCGCTTAATCGGATGTGAAGTGAAGGAGTGAACAAATCCCATGTGGACG 1235
 Db 1206 AACAGAGAGACTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1255
 QY 1236 GCTTCTAATCCCATGATAGCGGTCCATGATGACAGGCAATTTGGGGTAAATGACGTGA 1295
 Db 1266 GCTCTTCTGCGGCTGATCTTCAATGACATGATGATGATGATGATGATGATGATGAT 1325
 QY 1296 CGAATTCATGCTGCTGCTTTCGAGTGAAGTGGCGGCTGCTGCGCAACCCCGTTGG 1355
 Db 1326 CGAGTTCACCCAG 1385
 QY 1356 CTTCATACGCTTGGGCTCGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1415
 Db 1386 TTTCTTCTGCTTGGAG 1445
 QY 1416 GGGCAATTTGACACTGCTGTATATGATGATGATGATGATGATGATGATGATGATGAT 1475
 Db 1446 AGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1505
 QY 1476 TCGAGATGACCTACGCTGCTTATGATGATGATGATGATGATGATGATGATGATGAT 1535
 Db 1506 CACCCAGCGGCGGATGACCTGATGATGATGATGATGATGATGATGATGATGATGAT 1565
 QY 1536 C 1536
 Db 1566 C 1566

RESULT 3
 US-09-527-073-16
 ; Sequence 16, Application US/09527073
 ; Patent No. 6534313
 ; GENERAL INFORMATION:
 ; APPLICANT: Michael M. Neff
 ; APPLICANT: Joanne Chory
 ; TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
 ; TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING
 ; FILE REFERENCE: SALKINS_024A

Query Match 11.8%; Score 184; DB 4; Length 588;
 Best Local Similarity 99.0%; Pred. No. 1e-47;
 Matches 206; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
 US-09-527-073-16
 QY 1 ATGAG 60
 Db 382 ATGAG 441
 QY 61 AGT-CTTGTATATGAG 115
 Db 442 AGTCTGTATATGAG 501
 QY 120 ACATTTCTTAAACAG 179
 Db 502 ACATTTCTTAAACAG 561
 QY 180 AGAATCTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 207
 Db 562 AGAATCTTGT-AGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588

RESULT 4
 US-09-615-192A-101
 ; Sequence 101, Application US/09615192A
 ; Patent No. 6410718
 ; GENERAL INFORMATION:
 ; APPLICANT: Blockberg, Leonard N.
 ; APPLICANT: Havukkala, Ilkka
 ; TITLE OF INVENTION: Materials and Methods for the
 ; TITLE OF INVENTION: Modification of Plant Lipid Content
 ; FILE REFERENCE: 11000.10034U
 ; CURRENT APPLICATION NUMBER: US/09/615,192A
 ; CURRENT FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 08/975,316
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: US 08/713,000
 ; PRIOR FILING DATE: 1996-09-11
 ; PRIOR APPLICATION NUMBER: US 09/169,789
 ; PRIOR FILING DATE: 1998-10-09
 ; NUMBER OF SEQ ID NOS: 405
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 101
 ; LENGTH: 612
 ; TYPE: DNA
 ; ORGANISM: Eucalyptus grandis
 ; US-09-615-192A-101
 Query Match 4.8%; Score 75.8; DB 4; Length 612;
 Best Local Similarity 50.1%; Pred. No. 1.2e-13;
 Matches 216; Conservative 0; Mismatches 212; Indels 3; Gaps 1;
 QY 75 GAAGGATGCTCTCTGTATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 134

Db	185	GAGGGCGGTGAAC	TGGGTGTG	GGCTGAGGCCGCA	AAAGGCTGAGAGGCTTTGAGACAGCA	244
OY	135	AGGAATTCGAGGTCC	TCTTATCATTTCTTCATCGAAATGTTAA	GAACCTTGTAAT	194	
Db	245	AGGTCTTCGCCGCA	AGTCTCAACCTTCTGTGCGGCACTTCAAGGAAC	TGGGGAT	304	
OY	195	GATGCTTAAAGCTT	CTTCTCATCTTA	TGCTTTCTTCACAAATTTCTTCAGAGTTC	254	
Db	305	GCTCAAGGAAGCA	AGTCCAGCCCATGCGCTCTCGAGTACATCAAGCTGTGCTCTT	364		
OY	255	CTCTTTTACCATC	ACTGGAGAAATATCAGGTGCTACA	TTTCTGTTTGTTCGATCC	314	
Db	365	GCTTCTCTGCATCA	TCTTCCAAACCTA	TGGCAAGAC	TCGTTACATGATGGGCCC	424
OY	315	AACTTTCGGTTA	ACGTACCGGATCTCGATTTGATCAGAGATCTTCTTAAGTTCG	374		
Db	425	AACACCAAAAGTGA	ACATACGAACCCGGAACAAATAAAGAGGTATTTCTTAAGATATA	484		
OY	375	GTTCTACGAAAGAA	TAGAGCTCACCCCTTTGGTTAA	CAATTGAAAGCATGGACTACT	434	
Db	485	TGACTATCCCA	ACCACGCTCCATTCCTGTGTAAATGTGT---	CCGTATGATC	CGC	544
OY	435	TAGTCTCAAA	GTGAAAAATGGGCTCATCATCGAAAAATCATTTAGCCCTTACTTTCAAT	494		
Db	542	GAACTATGAGGGG	GAGAAATGGGCTCGGACCGAAGATTTATCAATCAGATTCACAT	604		
OY	495	GGAGAA	TTCTTA	505		
Db	602	GGAGAGTTGA	612			

```

RESULT 5
US-09-169-789-101
; Sequence 101, Application US/09169789
; Patent No. 665328
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c2
; CURRENT APPLICATION NUMBER: US/09/169,789
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 08/975,316
; EARLIER FILING DATE: 1997-11-21
; EARLIER APPLICATION NUMBER: US 08/713,000
; EARLIER FILING DATE: 1996-09-11
; NUMBER OF SEQ. ID NOS: 185
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 101
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-169-789-101

```

	Query Match	4.8%	Score 75.8	DB 4	Length 612
	Best Local Similarity	50.1%	Pred. No. 1.2e-11		
	Matches 216	Conservative	0	Mismatches 212	Indels 3
				Gaps	1
QY	75	GAAGGATATGTCCTGTTATGTGTGAGACCAAGAAAGATTGAAGACATTTCTCTAAATA	134		
DB	185	GAGGCGCGTAACTGGGTGTGGCTAGGCCCAAGAGGCTCGAGAGGTTCTGAGACAGA	244		
QY	135	AGGAATCGAGGTCCTCCTTATCATTTCTTCATCGSAAATTGTAAGAATTTGTTGAAT	194		
DB	245	AGGCTCTCCGGCAATCTCAACCTTCCTGGTGGGGAAGCTCAAGSAGAACTTGGGAT	304		
QY	195	GATGCTTAAAGCTTTCTTCATGCTATGCTCTTCTTCACAAATTTCTTCCGAGATTC	254		
DB	305	GCTCAAGAGGCCAAGCTCAAGCCATGCCGCTCCGATGACATCAAGGCTTGCTCTTT	364		
QY	255	CTCTTTTACCATCACTGAGAAAATCTACGGTGCTCAATTTCTGGTTGGTTCGATCC	314		

Db 365 GCCTTCTTGACATCATCTCTCCAAACCTATGCGAAAGACTCGTTACATGATGGGCCC 424

Qy 315 AACTTTCGGGTTAACGATAGCCGATCCGTGATTTGATCAGAGATCTTCTCTAAAGTCTGA 374

Db 425 AACCCAGAGTGAACATTTACGAAACCCGGAAACAATTAAAGAGGTATTTCTTAAGATATA 484

Qy 375 GTTTCACGAAAGAAATGAAGCTCACCCCTTGGTTAAACACTGGAAGCGCATGACACT 434

Db 485 TGAATATCCGAAGCCAGCCTCAATCCCTGGTAAAGTTGCT--CGCTATGAGCTCGC 541

Qy 435 TAGTCTCAAGGTGAAAAATGGGCTCATCATCGAAAAATCATTAAGCCCTACTTTTCATAT 494

Db 542 GAACCATGAGGGCGAGAAATGGGCTCGGCAACGAAAGATTATCATCATCAGCATTCACAT 601

Qy 495 GGAAGATCTTA 505

Db 602 GGAGAGGTGA 612

```

RESULT 6
US-09-615-192A-169
; Sequence 169, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows, Version 3.0
; SEQ ID NO 169
; LENGTH: 398
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-615-192A-169

```

Query Matchn	4.7%;	Score 73.4;	DB 4;	Length 398;
Best Local Similarity	51.2%;	Pred. No. 5,3e-13;		
Matches 198;	Conservative	0;	Mismatches 186;	Indels 3; Gaps 1
QY	563	ATAAGTTATCAAAAACGGTGAAGTTAGAGAGATGCTATGATGGTGTTCAGATTTTGA	622	
Db	6	ATTGTGATCGTAGAGAGATCCTGTGTGATCGACGGCTGGGTGCACCTTCAAAATTTGA	65	
QY	623	CTGAAGATGTTATTAGTAGAACAAGCTTTTGGAACTAGACTTAGAAGATGTGAGAGCTTT	682	
Db	66	CCCGTAGGTGATCTCTCGAACAGGCTTTGGCACTACCTTCGAAGAAAGCAAAAGATCT	125	
QY	683	TTGCACTTCAAGCTCAACAAATGCTTCTTTGTGCTGAAGCTTTCAAAAAGTTTATTC	742	
Db	126	CCGAACTTCAGGGGGAAACAAGCCAGCTCAAGATATATAGCCCTTCATCGGTTCATCTC	185	
QY	743	CTGGCTATAGATTTTTTCCGACAAAGGGAAATTTGAAGTCCGSAAGTTGACAAAGGAA	802	
Db	186	CTGGTTGAGGTTTGGCCACTAATGATGAACAGAGAGTAGAAGCATATGAATAAGAA--	243	
QY	803	TGAGGAAGTGGTGTGGAAGCTGATAGAGACGCGGAGACAAACGCTATATGATGAGAAAG	862	
Db	244	-AGTGGCGGCTCTGCTCTCATGAGACATATCCGCAAGAGAGAAAGCAATTAAGGGAAGGGG	302	
QY	863	GGGAAGAAATTAAGAGCCGCGCGGAAAGATTGTGGAGTTATATATTCAAGCAAAAG	922	
Db	303	AAGCTGCTGCGATGATCTGCTGGGCGTGTTCGAGATCAAAATGAAAGAAATGTGCG	362	

Mon May 24 08:10:08 2004

us-09-992-901-1.rn1

Page 6

```

; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-Fls
US-08-232-463-14

```

Query Match	3.4%;	Score 53;	DB 1;	Length 7218;
Best Local Similarity	9.3%;	Pred. No. 8.9e-06;		
Matches	23;	Conservative	137;	Mismatches 87; Indels 0; Gaps 0;

QY	762	GACAAAGGGAATTTGAAGTCGGAAAGTTAGAAAGAGAAATGAAGAACTGGTTTGA	821
Db	1262	RR	1203
QY	822	GCTGATAGACGGCGGAGACAAACGCTTATGATGAGAAAGGAGAAATGTAGAGACC	881
Db	1202	RR	1143
QY	882	GCGGGAAGGATTTGTTGGATTATGATTCAGGCAAGAAATGTACGTTCAAGACAT	941
Db	1142	RR	1083
QY	942	TGTGAGGAGGTAAAGCTTTTCTTGCCGGGAAAACAGACACTTCTATCTGCTGAC	1001
Db	1082	RRRRRRRRRRRRRRRATGCAAGCTCCCTGACCTGACCAAGCTCGGAATTAATCT	1023
QY	1002	GTGAGACG	1008
Db	1022	GTGAGCG	1016

```

RESULT 10
US-09-118-554-55
: Sequence 55, Application US/09118554A
: Patent No. 6365348
: GENERAL INFORMATION:
: APPLICANT: Reed, Steven G.
: APPLICANT: Xu, Jiangchun
: TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND
: TITLE OF INVENTION: METHODS FOR THEIR USE
: FILE REFERENCE: 210121.450C1
: CURRENT APPLICATION NUMBER: US/09/118,554A
: CURRENT FILING DATE: 1998-07-17
: EARLIER APPLICATION NUMBER: 08/968,255
: EARLIER FILING DATE: 1997-12-24
: NUMBER OF SEQ ID NOS: 67
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 55
: :
: LENGTH: 933
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-118-554-55

```

Query Match	3.1%;	Score 48;	DB 4;	Length 933;
Best Local Similarity	64.3%;	Pred. No. 9.1e-05;		
Matches	72;	Conservative	0;	Mismatches 40; Indels 0; Gaps 0;
OY	1342	AAACACCCGCTGGCTTCATACCGCTTGGCCCTCGAGTGGTGCATGTCAGATGTCAGAT	1407	
Db	329	ATACATCCCTAATCCCTTCATACCATCTTCACGTGATTAAGAACTGCATGGCGACGCT	388	
OY	1402	CTTGCTAATCTTCAGGCCCAATTGACACTCGCTGAATGATCCAAAGCTTCA	1453	
Db	389	TTGGCCATATATGATGTAAAGTGGAGTGGAGATTAACTGCTGCCCTTCA	440	

RESULT 11
US-09-118-627-55
; Sequence 55, Application US/09118627A
; Patent No. 6379951
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jianshun
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER

```

? TITLE OF INVENTION: AND METHODS FOR THEIR USE
? FILE REFERENCE: 210121.446C1
? CURRENT APPLICATION NUMBER: US/09/118,627A
? CURRENT FILING DATE: 1998-07-17
? NUMBER OF SEQ ID NOS: 67
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO. 55
? LENGTH: 933
? TYPE: DNA
? ORGANISM: Homo sapien
? US-09-118-627-55

```

Query Match	3.1%	Score 48	DB 4	Length 933
Best Local Similarity	64.3%	Pred. No. 9.1e-05		
Matches	72	Conservative	0	Mismatches 40; Indels 0; Gaps 0
Cy	1342	AAACAACCCCGTGGCTTCATCGCTTGAGCTCGAGATCGATACATGACATTGGTCAGAT	1401	
Db	329	ATACATCCCTATGCTTCATACATTTCTACGCTGGATTGAAGAACTGCATTGGGAGCAT	388	
Oy	1402	CTTGGTATACCTCAGGCCAATTACACTGCTGTATGATCAACAGCTTCA	1453	
Db	389	TTTGGCATATTGAGTGTAAAGTGGAGCTGGCATTAACCTCTGCTCGGCTTCA	440	

```

RESULT 12
US-09-602-877A-55
; Sequence 55, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiaochun
; APPLICANT: Dillon, David C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 55
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-55

```

```

Query Match 3.1%; Score 48; DB 4; Length 933;
Best Local Similarity 64.3%; Pred. No. 9.1e-05;
Matches 72; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 1342 AAACACCCCGGTGGCTCATACCGTTGGGCTCGAGATCGTACATGCTGTCAGAT 1401
DB 329 ATRCARTCCCTAAGCTTATACATCTTCTGACGTGATTAAGAGACGATTTGGCAGCAT 388
QY 1402 CTTCCTATATCTCAAGCCAAATTCACCTCGCTGTAAATGATCCACAGCTTCA 1453
DB 389 TTTCGCATATTTGATGTCTTAAGTGTGCAGATGCATTAATCTGTGTCGCTTCA 440

```

```

RESULT 13
US-09-602-877A-102
; Sequence 102, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107

```

Mon May 24 08:10:08 2004

US-09-992-901-1.rnt

Page 7

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 102
LENGTH: 940
TYPE: DNA
ORGANISM: Human
US-09-602-877A-102

Query Match 3.1%; Score 48; DB 4; Length 940;
Best Local Similarity 64.3%; Pred. No. 9.2e-05;
Matches 72; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

1342 AACACCCCGTGGTTCATCCGTTGGCTCGAGTGTGATCATGTCAGTAAT 1401
1330 ATACATCCCTATGCTTCAACCATTTCCAGCTGATTAAGAACTGCAATGGGCAAGAT 389
1402 CTGGTATACCTCAGGCCAATTGACACTCGCTGTAAATGATCCAAAGCTTCA 1453
390 TTGGCATATATTGAGTGAAGTGGAGTGGCATTAACTGCTCCGCTCA 441

RESULT 14
US-08-991-677-1
Sequence 1, Application US/08991677A
Patent No. 6252135
GENERAL INFORMATION:
APPLICANT: Chiang, Vincent L
APPLICANT: Carriaway, Daniel T
APPLICANT: Smeltzer, Richard H
TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
FILE REFERENCE: 50617
CURRENT APPLICATION NUMBER: US/08/991,677A
CURRENT FILING DATE: 1997-12-16
EARLIER APPLICATION NUMBER: US 60/033,381
EARLIER FILING DATE: 1996-12-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1708
TYPE: DNA
ORGANISM: Liquidambar styraciflua
FEATURE:
NAME/KEY: CDS
LOCATION: (48) ..(1571)
US-08-991-677-1

Query Match 3.0%; Score 46.2; DB 3; Length 1708;
Best Local Similarity 46.2%; Pred. No. 0.00049;
Matches 227; Conservative 0; Mismatches 258; Indels 6; Gaps 2;
1085 ATGTCCCTACCAAGACATGTCGTTAAAGCTTAAACGTTGATGATCTTGAACAGAT 1144
1057 GTGTCTGACCAAGATTTGACTTCAAGCTCCCTTATCTAATAGTGTACCAAGAGAG 1116
1145 CTTTAAAGTTGTATCCCAATAGTATGAT---TCGAGCGGCTAAATCGATATGA 1201
1117 CACTAAGGCTGACCTCCCAACACCACTAATGCTCCCTATGCGCCATGCAAGCTCA 1176
1202 AGCTAGAGGATCAAAATCCCATGTGGACGAGGCTTCAATCCCAATCAAGGATCC 1261
1177 AAATTGGTGTCTGACATCTTAAGGATCAAAATGTTATGATGATGCTGAGGCGGTG 1236
1262 ATATGACCAAGCAATTTGGGATATGAGTGAAGATTTCAATCACTGCTGTTGGG 1321
1237 CTCGTATCCAGCATGTGTGGC--TGACCACATAGATTTCGACGGAAGGATCTCTG 1293

1322 ATGAGTGGCGCGGTCGTCACCAACACCCGTTGCTTCAATACGTTGGCTCGAGATTC 1381
1294 AAGACGATGTGCATATAAAGGTGCAGATTATAGGCTACTGCGCTTGGTGGAGGAGGC 1353
1382 GTACATGATTTGTGACAAATCTGTATTAATCTTCAAGGCCAAATTTGACATCTGTATGA 1441
1354 GTGTTGGCCCGGTGACAACTTGGCATTAATTTGGTCACATCATGATAGGTCACCTAT 1413
1442 TCCAGCGTTC 1452
1414 TGACCATTTTC 1424

RESULT 15
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 2.9%; Score 45.8; DB 3; Length 4403765;
Best Local Similarity 53.0%; Pred. No. 0.097;
Matches 98; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
1357 TTCAATCCGTTGGCTGAGTTCGTATCATGATGATGATGATGATGATGATGATGATGAT 1416
1569612 TACCTGCCGCTTGGTGGGCGCGGACGATTCGATCGAGACGCTTCCCTGATGAG 1569553
1417 GCCAATTTGACACTGCTGTATGATCAACGCTTCACTTTCACTTGGCTCTTACTTAT 1476
1569552 AAGTTTGTAGGCGCGCATATATAGTCAACTTCACTTGACTCGGCCCTGGCTAT 1569493
1477 CAGCATGACCTACCGTCTTATGTTGTTATCTTCAATGATGATGATGATGATGATGATGAT 1536
1569492 CAGCTGAACTTCGAGCGGACGCTGACCTTGACCAAGCAGAGTACGATCATGCGG 1569433
1537 CGGAG 1541
1569432 AAGAG 1569428

Search completed: May 22, 2004, 20:15:52
Job time : 157 secs

Mon May 24 08:10:08 2004

us-09-992-901-1.rnpb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using BW model

Run on: May 22, 2004, 19:00:57 ; Search time 730 Seconds

(without alignments)
9730.698 Million cell updates/sec

Title: US-09-992-901-1

Perfect score: 1563

Sequence: 1 atgagagaagaagaatagcag.....tgaccatcatgagattga 1563

Scoring table: IDENTITY NJC

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database: Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	1563	US-09-992-901-1	Sequence 1, Appl 1
2	493.2	31.6	1236	US-10-424-599-59416	Sequence 59416, A
3	303	19.4	1778	US-10-424-599-40486	Sequence 40486, A
4	291.4	18.6	1844	US-10-424-599-56717	Sequence 56717, A
5	274.8	17.5	1886	US-10-425-114-30156	Sequence 30156, A
6	273.2	17.5	1836	US-10-425-114-22333	Sequence 22333, A
7	273.2	17.5	1877	US-10-425-114-13336	Sequence 13336, A
8	267.8	17.1	1853	US-10-425-114-1043	Sequence 1043, A
9	254.2	16.3	1922	US-10-424-599-85977	Sequence 85977, A
10	250.8	16.0	1872	US-10-424-599-126502	Sequence 126502, A
11	241.2	15.4	1557	US-09-938-842A-92	Sequence 92, Appl 1
12	241.2	15.4	1574	US-10-425-114-15634	Sequence 15634, A
13	232.4	14.9	1574	US-10-425-114-15634	Sequence 15634, A
14	231.4	14.8	4300	US-10-424-599-74806	Sequence 74806, A

15	230.2	14.7	1808	13	US-10-425-114-24276	Sequence 24276, A
16	218.6	14.0	1704	6	US-10-429-949-22	Sequence 22, Appl 1
17	194.8	12.5	1539	9	US-09-938-842A-379	Sequence 379, Appl 1
18	194.8	12.5	1539	11	US-09-938-842A-379	Sequence 379, Appl 1
19	191.2	12.2	1688	13	US-10-425-114-476	Sequence 476, Appl 1
20	185.6	11.9	1747	13	US-10-425-114-23110	Sequence 23110, A
21	184	11.8	588	9	US-09-992-901-1	Sequence 15, Appl 1
22	184	11.8	1776	13	US-10-425-114-13600	Sequence 13600, A
23	184	11.5	1367	13	US-10-424-599-104522	Sequence 26722, A
24	179.4	11.5	1539	13	US-10-424-599-104522	Sequence 104522, A
25	173.2	11.1	1539	13	US-10-425-114-28233	Sequence 28233, A
26	172.8	11.1	1126	13	US-10-425-114-22552	Sequence 22552, A
27	170.8	10.9	2188	13	US-10-424-599-76177	Sequence 76177, A
28	168.4	10.8	522	13	US-10-424-599-95836	Sequence 95836, A
29	160.2	10.2	1843	13	US-10-424-599-54624	Sequence 54624, A
30	153.6	9.8	1124	9	US-09-770-945-88	Sequence 88, Appl 1
31	147.6	9.4	1789	13	US-10-425-114-28468	Sequence 28468, A
32	146	9.3	698	13	US-10-424-599-116408	Sequence 116408, A
33	142.2	9.1	665	13	US-10-424-599-63982	Sequence 63982, A
34	132.2	8.5	345	13	US-10-424-599-95948	Sequence 95948, A
35	131	8.4	507	13	US-10-424-599-150803	Sequence 150803, A
36	126.2	8.1	1088	13	US-10-424-599-55315	Sequence 55315, A
37	121.4	7.8	1234	13	US-10-425-114-35022	Sequence 35022, A
38	115.4	7.4	748	13	US-10-425-114-3005	Sequence 3005, A
39	114.2	7.3	1491	13	US-10-424-599-107723	Sequence 107723, A
40	105.4	6.7	2147	13	US-10-424-599-76182	Sequence 76182, A
41	95.8	6.1	631	16	US-10-341-961A-234	Sequence 234, Appl 1
42	95.4	6.1	1904	13	US-10-425-114-22465	Sequence 22465, A
43	93	6.1	1793	13	US-10-424-599-47449	Sequence 47449, A
44	93	6.0	436	13	US-09-770-423-23	Sequence 23, Appl 1
45	89	5.7	1271	13	US-10-424-599-87597	Sequence 87597, A

ALIGNMENTS

RESULT 1
US-09-992-901-1
Sequence 1, Application US/0992901
Patent No. US20020073446A1
GENERAL INFORMATION:
APPLICANT: Neff, Michael M.
APPLICANT: Chery, Joanne
TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
FILE REFERENCE: MODULATED BRASSINOSTEROID SIGNALING
CURRENT FILING DATE: 2001-11-14
PRIOR FILING DATE: 1999-12-14
PRIOR APPLICATION NUMBER: US 60/170,931
PRIOR FILING DATE: 1999-03-16
PRIOR APPLICATION NUMBER: US 60/124570
PRIOR FILING DATE: 1999-03-16
PRIOR APPLICATION NUMBER: US 60/170,931
PRIOR FILING DATE: 1999-12-14
PRIOR APPLICATION NUMBER: US 60/172,832
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1: 1563
LENGTH: 1563
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-09-992-901-1
Query Match
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB
1 ATGAGAGAAGAAGAATAGCAGCTGTCATTCAGAAAGTTCTGTCTGTCTGAATCTTA 60
1 ATGAGAGAAGAAGAATAGCAGCTGTCATTCAGAAAGTTCTGTCTGTCTGAATCTTA 60

QY 61 AGCTGTAATAGTGAAGGATATCTCTGTATGTGTGAGACCAAGAAAGATTGAAGA 120
 DB AGCTGTAATAGTGAAGGATATCTCTGTATGTGTGAGACCAAGAAAGATTGAAGA 120
 QY 121 CATTTCTCTAAACAAGAAATTCAGAGTCTCTCTTATCATTTCTTCATCGGAATGTTAA 180
 DB CATTTCTCTCTAAACAAGAAATTCAGAGTCTCTCTTATCATTTCTTCATCGGAATGTTAA 180
 QY 181 GAACCTGTGTGAATGATGCTTAAAGCTTCTTCATCTATGCTCTTCATCAAAATTT 240
 DB GAACCTGTGTGAATGATGCTTAAAGCTTCTTCATCTATGCTCTTCATCAAAATTT 240
 QY 241 CTTCCTAGAGTCTCTCTCTTTTACCATCACTGAGAAATACTACGGTGTACATTTCTG 300
 DB CTTCCTAGAGTCTCTCTCTTTTACCATCACTGAGAAATACTACGGTGTACATTTCTG 300
 QY 301 GTTTGCTCGGTCCAACTTTCGGGTAAACGTTAGCCGATCTGATTTGATCAGAGATC 360
 DB GTTTGCTCGGTCCAACTTTCGGGTAAACGTTAGCCGATCTGATTTGATCAGAGATC 360
 QY 361 TTCTCTAAAGTCTGAGTCTTACGAGAAATGAAGCTCACTCTTGTGTTAACTTGA 420
 DB TTCTCTAAAGTCTGAGTCTTACGAGAAATGAAGCTCACTCTTGTGTTAACTTGA 420
 QY 421 GGGCATGACTACTAGTCTCAAGAGTGAAGAAATGGGCTCATCTGCAAAATCATTTAGC 480
 DB GGGCATGACTACTAGTCTCAAGAGTGAAGAAATGGGCTCATCTGCAAAATCATTTAGC 480
 QY 481 CCTACTTTTCATATGAGAAATCTTAAAGTCTGTACCAAGTGTGTAAAGAGTGTACT 540
 DB CCTACTTTTCATATGAGAAATCTTAAAGTCTGTGTACCAAGTGTGTAAAGAGTGTACT 540
 QY 541 GATATGCTGATAAATGATCGATTAAGTTATCGAAACGATGAAGTTGAGATGATC 600
 DB GATATGCTGATAAATGATCGATTAAGTTATCGAAACGATGAAGTTGAGATGATC 600
 QY 601 TATGAGTGTGTCAGATTTTGTGATGAAAGTGTATATGAGAAAGCTTTTGGAAATGAC 660
 DB TATGAGTGTGTCAGATTTTGTGATGAAAGTGTATATGAGAAAGCTTTTGGAAATGAC 660
 QY 661 TATGAGATGCTCGAGCAGTTTTCAGCTTCAAGTCAACAATGCTTTTGTGTGCA 720
 DB TATGAGATGCTCGAGCAGTTTTCAGCTTCAAGTCAACAATGCTTTTGTGTGCA 720
 QY 721 GCTTTTCAAAAAGCTTCACTCTGCTATAGTTTTTCCGACAAAGAGGAAATTTGAG 780
 DB GCTTTTCAAAAAGCTTCACTCTGCTATAGTTTTTCCGACAAAGAGGAAATTTGAG 780
 QY 781 TCTCGAAGTATGACAAAGAGATPAGAAAGTCTTTTGAAGCTGATPAGAGCGGAG 840
 DB TCTCGAAGTATGACAAAGAGATPAGAAAGTCTTTTGAAGCTGATPAGAGCGGAG 840
 QY 841 CAAGAAGCTATAGTGAAGAAAGGAGAAATGATPAGAGAGCGGCGGAGAAAGATTTGTG 900
 DB CAAGAAGCTATAGTGAAGAAAGGAGAAATGATPAGAGAGCGGCGGAGAAAGATTTGTG 900
 QY 901 GGATTAATGATTCAGGCAAGAAATGACGTTCAAGACATTTGTGAGAGAGTGTAAAGC 960
 DB GGATTAATGATTCAGGCAAGAAATGACGTTCAAGACATTTGTGAGAGAGTGTAAAGC 960
 QY 961 TTTTCTTCCGCGGAGAAACAGACATTTCTAATGCTGAGAGTGTGAGACACATCTGTGTA 1020
 DB TTTTCTTCCGCGGAGAAACAGACATTTCTAATGCTGAGAGTGTGAGACACATCTGTGTA 1020
 QY 1021 TCATGCAACCGGAGTGTGAGGCGCAAGACGATGAGATCTTCAAGGATCTTGGCGCTCA 1080
 DB TCATGCAACCGGAGTGTGAGGCGCAAGACGATGAGATCTTCAAGGATCTTGGCGCTCA 1080
 QY 1081 CGTATGTCCTTACCAAGGACATGTCGTTAAGCTTAAACGTTGATGATCTTGAAC 1140
 DB CGTATGTCCTTACCAAGGACATGTCGTTAAGCTTAAACGTTGATGATCTTGAAC 1140

QY 1141 GAGTCTTAAAGTGTATCCACCAATAGTACGATTCAGAGGCTAAATCGAGTGT 1200
 DB GAGTCTTAAAGTGTATCCACCAATAGTACGATTCAGAGGCTAAATCGAGTGT 1200
 QY 1201 AAGCTAGAGAGGTATCAAAATCCATGATGAGAGAGCTTCTAATCCCATCTATAGCGGTC 1260
 DB AAGCTAGAGAGGTATCAAAATCCATGATGAGAGAGCTTCTAATCCCATCTATAGCGGTC 1260
 QY 1261 CATCATGACCAAGCATTTTGGGGTAATGAGTGAACGAATTCATTCAGCTGCTTTCG 1320
 DB CATCATGACCAAGCATTTTGGGGTAATGAGTGAACGAATTCATTCAGCTGCTTTCG 1320
 QY 1321 GATGAGAGCGGCTGTGCGCAAAACCCCGTTGGTTTATACGTTGGCTCGAGATT 1380
 DB GATGAGAGCGGCTGTGCGCAAAACCCCGTTGGTTTATACGTTGGCTCGAGATT 1380
 QY 1381 CGTACATGCAATGTGTGAGAAATCTTGTATATCTTCAAGCCAAATTTGACATCGCTGTAATG 1440
 DB CGTACATGCAATGTGTGAGAAATCTTGTATATCTTCAAGCCAAATTTGACATCGCTGTAATG 1440
 QY 1441 ATCCAGCGCTTCACTTCTCACTTGTGCTCTTATACAGACATGACATCGCTGTAATG 1500
 DB ATCCAGCGCTTCACTTCTCACTTGTGCTCTTATACAGACATGACATCGCTGTAATG 1500
 QY 1501 TTGCTTATCTTCAACATGATGTGACCAATCACCTTCCGAGATTGACCAATCATGAGAT 1560
 DB TTGCTTATCTTCAACATGATGTGACCAATCACCTTCCGAGATTGACCAATCATGAGAT 1560
 QY 1561 TGA 1563
 DB 1561 TGA 1563

RESULT 2
 US-10-424-599-59416
 ; Sequence 59416, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 59416
 ; LENGTH: 1236
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1) (1236)
 ; OTHER INFORMATION: unsure at all n locations
 ; FEATURES:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_24664C.1
 US-10-424-599-59416

Query Match 31.6%; Score 493.2; DB 13; Length 1236;
 Best Local Similarity 70.1%; Pred. No. 2,1e-142;
 Matches 733; Consilative 0; Mismatches 283; Indels 30; Gaps 4;

QY 21 CTGCTTCAATTCAGAAAGTTCTTGTCTGTGATCTTAAATCTTGAATGAGAGG 80
 DB CTGCTTCAATTCAGAAAGTTCTTGTCTGTGATCTTAAATCTTGAATGAGAGG 80
 QY 189 CTTCCTCTCCCTCAAGCTTCTTCTTCTTCACTTCACTGCTCTTCTCTCTCAAGCT 248
 DB CTTCCTCTCCCTCAAGCTTCTTCTTCTTCACTTCACTGCTCTTCTCTCTCAAGCT 248
 QY 81 TATGCTCTGTTATGAGGAGACCAAGAAAGATTGAAGAACTTCTTAAACAAGAAAT 140
 DB TATGCTCTGTTATGAGGAGACCAAGAAAGATTGAAGAACTTCTTAAACAAGAAAT 140
 QY 249 CACGCTGTGCTGTGGGAGGACCAAGAAATGAGGCACTTCTCCAAAGCAAGGAT 308
 DB CACGCTGTGCTGTGGGAGGACCAAGAAATGAGGCACTTCTCCAAAGCAAGGAT 308
 QY 141 TCAGAGTCTCTTATATCTTCTTCAATCGGAAATGTTAAAGAACTTGTGGAATGATGCT 200

Mon May 24 08:10:08 2004

us-09-992-901-1.rnpb

Page 3

Db	309	CAGAGCCCCCTTATCTTCTTCAATGGAAACGTTAAGAACTTGTCCGAGATGTT	368
Qy	201	TAAAGCTTCTCTCACTCTTAAGCTTCTCTCAAAATCTCTTAAGTCTCTCTTT	260
Db	369	GAAGGCTCTTGTCAACTTAAGCTTCTCTCCATAACATCTTCCCAAGTCTTCTCTT	428
Qy	261	TTAACCATCATCTGAGAAATACTACGCTGCTACATTTGCTTTGGTTCCGTCACATTT	320
Db	429	CTACCACTCACTGAAAGAAATTTATATGAGCAACATCTCTAGTTGGTTGACCACTGT	488
Qy	321	CCGGTTAACGCTAGCCGATCTCGATTGTGATCAGAGATCTTC---TCTAAGTCTAGTT	377
Db	489	TGCTCTCACCGTCTCTGAGCCAGACCTCATCCGGAAATCTTCAATCCAGTCAAGATT	548
Qy	378	CTACGAAAGAAATGAAGCTCACCCCTTGGTTAAACAATTGAAGGCATGACTCTTAG	437
Db	549	TTACGAGAAAAACGAAGCTCCACACACTGTGGAAGCAGCTAGAAGGTATGGGCTCTCTAG	608
Qy	438	TCTCAAAAGTGAATAATGGGCTCATCATCGAAAAATCACTAAGCCACTTTCAATAGA	497
Db	609	CTCCAAAAGAGAAAAATGGGCTCCACACAGAAAAATTTATCTCTCCACTTTTCACATGA	666
Qy	498	GAATCTTAAGTGTCTGTGTACAGATTGTGTGAAAGGTGACTGATATGTGATGAATAATG	557
Db	669	AAATCTCAATGTCTGTGTACCGGTGATGCAACAAGGTGTGAGATCTGTGAATAATG	728
Qy	558	GTCGCAATAACTTATCAGAAAACGTTGAAGTTGAGTGTGATGTCTATAGATGGTTCAAT	617
Db	729	GTGGGCAA---TGGGTGAGAAAGGCGCAGATGAATGAAGTTCCGAATGGTTCCAAG	785
Qy	618	TTTGACTGAAGATGTATTATATAGAAACAGCTTTTGAAGTACTATGAAGTGTCTGAGC	677
Db	786	CTTAAAGAAAGATCATTTACGAGGACCGCGTTTGAAGACGCTACGAAGATGGCAAGC	845
Qy	678	AGTTTTTGCACCTCAAGCTCAACAAATGCTTCTTTGTCTGAAGCTTTCCAAAAAGCTTT	737
Db	846	CATTTTCCGATTACAGGCCCAACAATGATGACTTGGCCGCGATGCTTCCAAAAAGTATT	905
Qy	738	CATTCTGGCTATAGATTTTTCGACAAAGGGAATTTGAAGTCTCGAGTGTGAACA	797
Db	906	CATCCAGGTATAGATTTCTTCCACAGAGGAAATATAGATCTTGGAAATTTGGAGAA	965
Qy	798	GGAGATTAAGAAAGTCTTTGTTGAACTGATAGACGCGCGGAGCAAAACGCTATATGATGG	857
Db	966	GGAGATTAAAGAAATGCTGCTGATGATCTATCTCAAGCGAGAGAAATGAGAAAGGGTG	1025
Qy	858	AGAAAGGGAAGAAATGTAAGAGACCGCGCGCAAGATTTGTGGGATTTATATCAATCGGC	917
Db	1026	TGGGTTGAGGAA---AAAGAAAAAGGCGCAAGGATTTGTGGAGCTATGATTTTCAGGC	1082
Qy	918	AAAG-----ATATGACGGTTACAGACATTTGTGAGAGAGATGTA	956
Db	1083	ATCGAATATAATATAATATGTCATATGTCACTAGTGAATGACATGGTGGAGAGTGCAA	1142
Qy	957	AAGCTTTTCTTGGCCGGGAAACAGACAATTTCTATATGTGTGACGTGACGACATCTT	1016
Db	1143	GAGCTTTTCTTGTGACGCGCAAAACAGACACATCCACCTGTGACGTGACGACATCTCT	1202
Qy	1017	GCTATCCATGACCCCGAGTGGCAGG	1042
Db	1203	CTTGTGTATGACCCACACTGGCAGG	1228

```

1  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
2  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
3  FILE REFERENCE: 38-21(53223)B
4  CURRENT APPLICATION NUMBER: US/10/424,599
5  CURRENT FILING DATE: 2003-04-28
6  NUMBER OF SEQ. ID NOS: 265684
7  SEQ. ID NO 40486
8  LENGTH: 1778
9  TYPE: DNA
10 ORGANISM: Glycine max
11 FEATURE:
12 OTHER INFORMATION: Clone ID: PAT_MRI3847_136559C.1
13 US-10-424-599-40486

```

Query Match	19.4%	Score 303;	DB 13;	Length 1778;
Best Local Similarity	54.1%;	Pred. No. 6.6e-83;		
Matches 719;	Conservative 0;	Mismatches 585;	Indels 25;	Gaps 4;

[illegible]

QY 1052 GTGATGAGGCTCTCAGAGTCTGGGCTCAGTGTATGCTCCCTACCAAGAGCATGCTTAA 1111
DB 1088 GCAAGAGATCTCTCATGTCATGAGCGAATCGACTTCTGCTGACAGACATTTAAATG 1147
QY 1112 AGCTTAAACGTTAGATGATCTTGAAGAGCTTTTAAAGTTGATCCACCATATGAG 1171
DB 1148 ACCTTAAGATTGTACCATGATATTAAGAAACATTCGGCTCATCTCCAGACATTA 1207
QY 1172 CTACGATTGACGCGCTTAATCGGATGTGAAGCTAGAGGCTACAAAATCCCATGTGGCA 1231
DB 1208 TGTATATGAGGCAAGCTCTAAAGATGTATGCTTGGAAAGCTTAATATCTGCTAAGA 1267
QY 1232 CGGAGCTTCTAATCCCATCATAGCGGTCCATCATGACCAAGCCATTTGGGGTAAATGACG 1291
DB 1268 CTCAACTCTTTTGGGATGAGCTGTGTCATCATGACAGAGAAATCTGGGGAGAAAGATT 1327
QY 1292 TGAACGATTCATTCAGAGCTCGGTTGGGATGAGATGGCCGCTGCTGCCAAACACCCCG 1351
DB 1328 ACCTTAACCTTCAATCCGATGAGATTAGCCACGGCT-----CCGCCACGCG 1375
QY 1352 TTGGCTTCATACCGTTTGGCTCGAGATTCTGATCATGATGTGTCAGATCTTGCTATAC 1411
DB 1376 TCGATTCCTTCTCCCTTCGATTAAGCCCTCGAATCTGTGTGGGCAAAATTTGGCATTGG 1435
QY 1412 TTGAGGCCAAATTTGACACTGCTGTATATGATCCAAAGCTTTCACCTTTCACTTGGCTCTA 1471
DB 1436 TTGAGGCCAAAGATTGCTTGTGCTTGTATCATCAAGTTTGTGCTTATCTCCAA 1495
QY 1472 CTATCAGCATGACCTACCGTCTTATGTTGCTTATCTTATCCATGATGTCACCAATCA 1531
DB 1496 ACTATATGATGCTCTCTATATCTATTTGTATCTTGGAGCTCTCAATATGTCACAAATCA 1555
QY 1532 CCTTCCGGA 1540
DB 1556 TTTTCAGAA 1564

RESULT 4
US-10-424-599-56717
Sequence 56717, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424, 599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ. ID NOS: 285684
SEQ. ID NO 56717
LENGTH: 1844
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_22226C.1
US-10-424-599-56717

Query Match 18.6%; Score 291.4; DB 13; Length 1844;
Best Local Similarity 51.4%; Pied. No. 2.8e-79;
Matches 763; Conservative 0; Mismatches 696; Indels 24; Gaps 3;

QY 203 AAGCTTCTTCTCATCTATGCGCTTCTCTCAAAATATCTTCTTAAAGTTCTCTTTT 262
DB 383 AAGCAAGATCCAAACCATGATCTCTCCATGATATGTGCGACGTGTGTCACATC 442
QY 263 ACCATCATGAGAAATCTAAGTGTCTATATCTGTTGGTTGGTGGCAACTTCC 322
DB 443 TGCATCGAAGTCTCTCAACATGCGAAGATTTCTTATTTGTTGGACCAACCA 502
QY 323 GGTAAAGGTAGCCCATCTGATTTGATCAGAGATCTTCTTAAGTCTGAGTTACG 382
DB 503 GGGTACCCCTCAAGATCTGAGCTGATCAAAAGTATCTTAACAGATCAGTATTTCC 562
QY 383 AGAAGATGAGACTCACCTTTGTTAAACAATTGAAGGAGATGACTACTAGTCTCA 442
DB 563 GAAAGCTGAAACCAATCAGCTTGGCAATTAATCT--AGTACTGTTCTGTAACTATG 619
QY 443 AAGTGAATAATGAGCTCATCATCGAAATATCATTAGCCCTACTTTTCAATGAGATC 502
DB 620 ATGAGAAATAATGAAACAAGACAGAAATTAATATCTGATTCAGTTTGAATAAT 579
QY 503 TTAAGTCTGTGACAGTGTGTGTTAAAGTGTGACTGATATGATGATTAATGTCG 562
DB 680 TGAAGATTATGTTACCAATATCTTCAAAAGTTCATGATCTAATTAATCAAGTGAAG 739
QY 563 ATTAAGTTATCAAAAAGGTGAAGTGAAGTGTATGATGATGTTGAGATTTGA 622
DB 740 GAATGTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 799
QY 623 CTGAAGATGTTATGATGAACAAGCTTTTGAAGTATGATGATGATGATGATGATGATGAT 682
DB 800 CTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 859
QY 683 TTGATCTCAAGCTCAACAAATGCTCTTGTGCTGAAGCTTTCAAAAGTCTTCAATC 742
DB 860 TTCACTTCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 919
QY 743 CTGCTATGATTTTTCGACAAAGGGAATTTGAAGTCTGGAAGTTACAGAGAGA 802
DB 920 CTGATGAGATTTTGTACTCTGACCAATCGAGATGAGAAATGATGATGATGATGATGATGAT 979
QY 803 TAAAGATC-----GTTGTTGAAGCTGATGAGCGCGGAGCAAAACGCTATG 853
DB 980 TAAAGCTCACTTACAGACATGATTAAGAAAGAGAGAAACGCAAAAGCAGGTAAAG 1039
QY 854 ATGAGAAAGGAGAAATGATGAAGAGCGGCGGAGAAAGATTTGTGGATTAATGATTC 913
DB 1040 CTACTAGGATGATCTTGTAGTATCTTGTGAGTCAAAATCAAGAAATTAACAGAAC 1099
QY 914 AAGCAAGATGACGCTC-----AGCATTTGTGAGAGAGTAAAGCT 961
DB 1100 AAGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1159
QY 962 TTTTCTGCGCGGAAACAGACAACTTCTAATCTGCTGACGTGACGACATCTTGTAT 1021
DB 1160 TCTACTTTCAGAGGAGAGACCACTTCACTTCTGTTGACATAGTGTGTTAA 1219
QY 1022 CCATGACCGGAGTGGCAGCGCAAAAGCAGTATGATGATGATGATGATGATGATGATGATGAT 1081
DB 1220 GTAGTACCTTATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1279
QY 1082 GTGATGCTCCCTACCAAGACATGCTTGAAGCTTAAACGTTAGATGATGATGATGATGATGATGATGAT 1141
DB 1280 AAGCAACAAATTTGATGAGGCTTAAGTCACTTAAATGTCACCATGATTTTGTGAAG 1339
QY 1142 AGCTTTAAGTGTGATCACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1201
DB 1340 TGGTCTTGAAGCTATACCAACGAGCTTGTCTTATGAATGTTGACAGATATGA 1339
QY 1202 AGTGAAGAGGTACAAATCCCATGTCGACGAGCTTCTAATCCCATGATGATGATGATGATGATGATGATGATGAT 1261
DB 1400 AACTGGAAGACTATCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
QY 1262 ATCAGTACCAACCATTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1321

Mon May 24 08:10:08 2004

us-09-992-901-1.rmpb

Page 6

Db 1662 GGTGCAAAATTTAGCTGAAGAAAT 1667

RESULT 6

US-10-425-114-22323
; Sequence 22323, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22323
; LENGTH: 1836
; TYPE: DNA
; ORGANISM: Oryza sativa japonbare
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3431-054-B3_FLI
US-10-425-114-22323

Query Match 17.5%; Score 273.2; DB 13; Length 1836;
Best Local Similarity 51.2%; Pred. No. 1,4e-73;
Matches 750; Conservative 0; Mismatches 698; Indels 18; Gaps 4;

QY 94 TGGTGAAGACCAAGAAATGAGAAATCTTCTCAACAGAGATTGAGTCTCT 153
DB TGGTGAAGACCGGCGGCTGAGCGCGCCCTCGGGCGAGGAGATCCGGGCAACCG 239
QY 154 TATCATTTCTTCATCGAAATGTTAAAGAACTGTGGAATGATGCTTAAAGCTTCT 213
DB TACCCGCTTCACCGGAGAGTCCGAGAGAGTCCGGCTCAACCGGAGCGCGAGAG 299
QY 214 CATCTATGACCTT---CTCTCAATATCTTCTAGAGTCTCTCTTTACATCAC 270
DB AAGCGCTGCGGCTGCGTCCGACGATCATCCGCGGCTGCGCATGTCGCAAA 359
QY 271 TGGAGAAAATCTACGAGTCTACATTTCTGTTTGGTTCGTCAACTTTCGGTTAA 330
DB GCCGTGAGAGACAGGGAACCATCTTCACTTGGTTGGCCCAAGCGCAAGTGA 419
QY 331 GTAGCGCATCTGATTTGATGAGAGATCTCTCAAGTCTGAGTTCTAGAGAAAT 390
DB ATTGAGACCTGAAATCAATAGGAAATTAATGTTATGTTTGGCCACTATGCAAA 479
QY 391 GAAGCTACCTTTGGTTAAACAATTGAAGGAGATGAGACTTATGTCGAAGTGA 450
DB CCAAGCTTACCGGCTGGAAGATGCTAGCTCCGAGTTGTAAGCTATGAGGCGAG 539
QY 451 AAATGGCTCATTCGAAATCATTAAGCCATCTTTCAATGAGAAATCTTAAAGTTG 510
DB AATGGGCAAGACCGGAGAAATCTGATCTGCTTTCACAGAGAAATTAAGCGG 599
QY 511 CTGTGACAGTTGTGTGAAGATGTGATCTGATATGTTGAGTAATGTCCATAATGA 570
DB ATGCTGCAAGTTTCTTAATCTGTCACGAAATGTTTACAGAGTGAAGATTCATG 659
QY 571 TGAGAAACGAGTGAAGTGAAGTATGATGAGAGGTTTCAATTTTGAAGTAAAT 630
DB TCTATTTGAAGAAATGTCAGAGTGAATGTTTGGCTGAGTTCAGAAATTTTCAAGAGAT 719
QY 631 GTTATTTAGTAAACAGCTTTTGAAGAGTGAAGTGAAGTGTGAGCAATTTTGAAT 690
DB GTCATATCAAGAGACGATTTGGTGAAGCTATGAGAGAGAGAAATTTTTCAGCTG 779

QY 691 CAAGCTCAACAAATGCTTCTTTGCTGAAGCTTTTCAAAAAGCTTCAATCTGAGTAT 750
DB CAGCAGAGTGCAGCGGAACGATATATACAGCTTTTGGCAATTTTATACAGATAT 839
QY 751 AATTTTTCGACAAAGAGAAATTTGAATCTTCGAAATTTAGACAGAGATTAAGAA- 809
DB TGGTTCTTACCAACTTAAACAAAGAGAGTGAAGAAATTTGAAAGAGAGTCAAGAA 899
QY 810 -----CTGCTTTTGAAGCTGATGAGAGCGGAGACAAACGCTATGATGAGAA 861
DB CTCTACAGAGAAATATTTGAAAGAGAGAGCGGCTATTTAAATATGTTGAACAGTAAT 959
QY 862 GGGAGAAATGTTAGAGAGCGCGGAGAGATTTGTTGGATTAATGATTCAGCAAG 921
DB GGTGACTTGTGGCTTATTTGTTGAGTCAAAATATGAGGAGTCAAAATGGAAAGCAGAA 1019
QY 922 AATG---TGAGCGTTCAAGACATTTGAGAGAGTAAAGCTTTTCTCGCGGAGAA 978
DB CTAGAGATGATCTAGAGAGAAATTTTGAAGATGACAGTATTTTATTTTGGAGAAAG 1079
QY 979 CAGACACTTCTAATCTGCTGAGTGAAGACATCTTCTATTCATGACCCGAGTGG 1038
DB GAGACAACTCAGTATTTGCTCACTTGAATTTATTTGCTAATGATGACCCGAGTGG 1139
QY 1039 CAGGCAAGAGACAGTATGAGAGTCTCAAGGCTCGCGGCTCAGGTATGCTTACCAAG 1098
DB CAGAGCGAGCAAGAGAAAGTCTTACACCATTTGG---AGAAACACACAGACTAT 1196
QY 1159 GACCATGCTGTTAACTTAAACGTTGAGTATGATCTTGAACAGAGCTTTAAAGTTGAT 1158
DB GATGCTTAAAGCTGCTGATATGATGATGATGATGATGATGATGATGATGATGAT 1256
QY 1257 CCGCAGTGTGTTTCTTGAACAGAGAACTTCAACCATTTGG---AGAAACACACAGACTAT 1316
DB ATCCAGTGTGACAGAGCTTCTAATCCATCAATGAGTCTCAATGACAGCAAT 1278
QY 1317 TATCCGCTGAGAGTATACCTTATGTTGCTTATTTATTTATTCATGATTCGCAAT 1376
QY 1279 TGGGTAATGAGAGTGAAGATTCATTCAGCTGCTGCTGAGTGAAGGCGGCTGCT 1338
DB TGGGTAATGAGAGTGAAGATTCATTCAGCTGCTGCTGAGTGAAGGCGGCTGCT 1436
QY 1339 GCCAAACACCCGTTGCTTCAATACGTTTGGCTCGAGTGTGATATGATTTGCTG 1398
DB ACGAAGTATCAGACCTCTTCTTCCATTTGGATGGGCTCCCGAATCTGCAATCGGCGAG 1496
QY 1399 AATCTGCTATCTTCAAGCGCAATGACACTGCTTATGATTCAGAGCTTCACTT 1458
DB AATCTGCTATCTTCAAGCGCAATGACACTGCTTATGATTCAGAGCTTCACTT 1556
QY 1459 CACTTGGCTCTTCACTTCAAGTATGACACTGCTTATGTTGCTTATTCACAT 1518
DB GAGCTTCAACATGATATTCAGAGCACTGATGATGATGATGATGATGATGATGAT 1616
QY 1519 GGTGACCAATCACTTCCGAGAT 1544
DB GGTGACCAATTTAGCTGAAGAAAT 1642

RESULT 7

US-10-425-114-13396
; Sequence 13396, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 13396
 ; LENGTH: 1877
 ; TYPE: DNA
 ; ORGANISM: *Oryza sativa japonbare*
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: JC-05LELIB3474019601_FLI
 US-10-425-114-13396

Query Match 17.5%; Score 273.2; DB 13; Length 1877;
 Best Local Similarity 51.2%; Pred. No. 1.4e-73;
 Matches 750; Conservative 0; Mismatches 698; Indels 18; Gaps 4;

QY 94 TGGTGAAGCAAGAAAGATTGAAGACATTTCTTAAACAAGGAATTCGAGTCCCT 153
 DB TGGTGAAGCAAGCAAGAAAGATTGAAGACATTTCTTAAACAAGGAATTCGAGTCCCT 153
 QY 154 TATCATTTCTTCATCGAATGTAAAGAACTTGTGAATGATGCTTAAAGCTTCTTCT 213
 DB TACCGCTCTTACCGGCGAGCGTCCGGAAGCTCCGCTCAACCGGAGGCCCGGAA 338
 QY 214 CATCTAAGCTTT---CTCTCAATATTTCTTCTAAGTTCCTTTTTCATCAC 270
 DB AAGCGCTGCGCGCTGCGCTGCGACATCAATCCCGCGCTGCGCTGCGCTGCGCTGCGCT 338
 QY 271 TGGAGAAATCTACGAGTCAATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 330
 DB GCGCTGAGGAGCAAGGAAACATCACTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 458
 QY 331 GTAGCCGATCTGATTTGATCAGAGATCTTCTCTAAGTGTGATGTTCTACGAGAAAT 390
 DB AATTGACACCTGAAATCAATAGGAAATGTTATGTTCTAATGTTGCTCACTATGCGCAA 518
 QY 391 GAAGCTCACTCTTGTGTTAAACAATTGAAGCGATGAGTACTTAACTTCAAGGTGA 450
 DB CCAAAGCTTACCCGCTCGGAAGTTGCTAGCTCCGAGTTGTAAGTAAAGGCGAG 578
 QY 451 AATGCGCTCATATCGAATAATCAATAGCCCTACTTTCTAATGAGATCTTAAATG 510
 DB AATGCGCTCATATCGAATAATCAATAGCCCTACTTTCTAATGAGATCTTAAATG 510
 QY 511 CTGTGACGATGTTGTGTAAGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 570
 DB ATGCTGCGAGTTTCTTACTGCTGACGGAATGTTGTTCAAGATGAGAAATTTCAATG 638
 QY 571 TCAGAAACGATGAAGTGAAGTGAATGTTGATGAGTTCAGATTTGACTGAAGAT 630
 DB TCAATGAAGATGCAAGATGATGTTGCTGAGTTCCAAATCTTACAGAGAT 758
 QY 631 GTTATAGTGAAGCGCTTTGAGAGTCAATGAGATGAGTGAAGTGAAGTGAAGTGAAGT 690
 DB GTCAATGCAAGCGATTCGTAAGCGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 818
 QY 691 CAAGCTCAACAATGCTTCTTGTGCTGAAGCTTTTCAAAAAGTCTTCAATCTGAGTAT 750
 DB CAAGCTCAACAATGCTTCTTGTGCTGAAGCTTTTCAAAAAGTCTTCAATCTGAGTAT 750
 QY 751 AGATTTTTCGCAAGAGGAAATTTGAAGTCTCGAAATGTAAGCAAGATTAAGAA- 809
 DB TGGTCTTACCAACTTAAACCAAGAGTGAAGAAATTTGAAGAGAGTGAAGAA 938
 QY 810 -----GTCGTTGTTGAAGCTGATAGAGCGCGGAGCAAAACGCTATAGTGAAGAA 861
 DB CTCTTACGAGAAATATTGAAAGAGAGCGCGGCTATTAAATAATGTAAGAAACGAT 998
 QY 862 GGGGAAAGATTAAGAGCGCGCGGAGAAATTTGTTGGATTTAATTTACAGCAAG 921
 DB GGTGACTGTTGTTGGCTTATTGTTGAGTCAAAATATGAGGAGTCAAAATGGAAGCAAGAA 1058

QY 922 AATG---TGACGTTCAAGACATTTGAGAGAGTGTAAAGCTTTTCTTCCCGGAAA 978
 DB CTAGAAATGACTACGAGCGAAATTAATGAGAAATGAAAGTATTATTTATTTGAGAAAG 1118
 QY 979 CAGACAACTTCTAATCTGCTGAGCGAGCAACATCTTGTCTATCCATGACCCGAGTGG 1038
 DB GAGACAACTTCTAATCTGCTGAGCGAGCAACATCTTGTCTATCCATGACCCGAGTGG 1178
 QY 1039 CAGCCAAAGCACTGATGAGAGTCTTGAAGGTCTGAGGCTCAGCTGATGCTTACCAAG 1098
 DB CAAGGCGAGCAAGCAAGAGAGTCTTGAAGGTCTTGAAGGTCTTGAAGGTCTTGAAGGT 1235
 QY 1099 GACCATGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 1158
 DB GATAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAA 1295
 QY 1159 CCACCAATAGTACGATTCGAGCTGCTTAAATCGATGTAAGCTAGAGGCTACAA 1218
 DB CCGCAGAGTGTGTTCTTGAACGAGCAATACAGGAAATGAGGCTCGCGGCAATCAA 1355
 QY 1219 ATCCATGTCGAGAGCTTCTATCCATCATAGCGGTCCATCATGACCAAGCAT 1278
 DB TATCCGCTGAGAGTACCTTATGTTGCTTATTTATTTATTTATTTATTTATTTATTTATTT 1415
 QY 1279 TGGGTATGAGCTGAAGCAATTCATCAAGCTGCTTGGTGGATGAGTGGCGGTGCT 1338
 DB TGGGAAAGATGAGAGTGAATTCATTCAGAGAGGTTGCTGATGCTATCTTCAACGCA 1475
 QY 1339 GCGAAACCGCGTGTGCTTATCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1398
 DB ACGAAGTATCAGACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1535
 QY 1476 AATCTGTACTACTCAGCGCAATGACACTGCTGTAATGATTCACAGCTTACCTTT 1458
 DB AACTTGTACTACTCAGCGCAATGACACTGCTGTAATGATTCACAGCTTACCTTT 1595
 QY 1459 CACTTGTCTTACTTATCAGCATGACCTTACCTTCTTATGTTGCTTATCTTCAACAT 1518
 DB GAGCTTGTCTTACTTATCAGCATGACCTTACCTTCTTATGTTGCTTATCTTCAACAT 1655
 QY 1519 GGTGACCAATCACTTCCGAGAT 1544
 DB GGTGACCAATCACTTCCGAGAT 1681

RESULT 8
 US-10-425-114-1043
 ; Sequence 1043; Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaka, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 1043
 ; LENGTH: 1853
 ; TYPE: DNA
 ; ORGANISM: *Zea mays*
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700100628_FLI
 US-10-425-114-1043

Query Match 17.1%; Score 267.8; DB 13; Length 1853;
 Best Local Similarity 51.4%; Pred. No. 6.6e-72;
 Matches 679; Conservative 0; Mismatches 627; Indels 15; Gaps 2;

OY	228	CTCTCAAAATATCTTCTCTAGAGTCTCTCTTTTAAACAATCACTGGAGAGAAAATCTACCG	287
Db	310	CTGCCAGACATCAACCCACGCGGACGCCCATGATCAAGCAACCATCAAGAAATACGG	369
OY	288	TGCTACATTTCTGGTTGGTTCCGGTCAACTTCCGGTAAACGGTACCGATCTCTGATTT	347
Db	370	GAACCTATTCGTTCACTCGTTCCGCCCAACACCAAGGGGTGATATTTCTACCAAGATTT	429
OY	348	GATCAGAGAGATCTCTCTATAGTCTGAGTTCTACGAAAGATGAAGCTCAACCTTTGGT	407
Db	430	AGTCAAAGAGGTGCTGTCTAATAAGTTTGGCCACTTTGGCAAAACCAAGAGTACCGCAT	489
OY	408	TAAACAACCTGAAGCGCATGCACTACTAGTCTCAAAAGGTGAAAATATGGGCTCATATCG	467
Db	490	TGGAGAGTTGCTACGCCAACCGGCTGTGTAATCATGATGGTGAAAATATGGCAAAAGCACG	549
OY	468	AAAATCATATAGCCCTACTTTTCAATATGAGAAATCTTAAAGTGGCTTGTACAGATGTGTT	527
Db	550	GAGAAATCTTATCTCTGATTTCAACAATGAAAATATTAAGGAGATGATGCAAGATTTTC	609
OY	528	GAAAGATGTGACTGATATGTGTGATTAATGTGCTCCGATTAAGTTATCAGAAAAACGTGAAGT	587
Db	610	TACCTGCTGTATTTGAAATGATTAAGTATGAGATGAATTAATCAATGTCTTCCAGGAGATCTTC	669
OY	588	TGAGGTAGATGCTATGATGAGTGGTTCAAGATTTTCACTGAAGATGTATATAGTAGAACGC	647
Db	670	TGAGATAGATCTTGCGCTGAGTTCCAGATCTTACTGGAGATGTATCTTCAGAACTGC	729
OY	648	TTTTGAAAGTACTATGAAAGATGCTCGACAGCTTTTTCAGCTTCAACTCACAAAAGCT	707
Db	730	GTTTGGGAGCAACTATCAAGAGGAGAGGAAATTTTGAAGTCAAGAGAACTAGCTGA	789
OY	708	TCTTTGCTGAAGCTTTCAAAAAGCTTATCTCCGCTATGATTTTTCGACAG	767
Db	790	ACGCTCATCCAAATCTGTTCAACAATTTTATCCAGGCTATTTGGTTCTTGGCCACAA	849
OY	768	AGGGAATTTGAATCTCGGAAGTTGACAGAGTTAAGAAATGCAAGTCTGTGAAAGCTGAT	827
Db	850	AAACAAACGAAGATGAGAGCAATCGATGAGATTCGCAAAATTCCTCGGTGAATTAAT	909
OY	828	AGAGCGCGGAGACAAACCGCTATGATGAGAAAGGAGAGAAATGTAAGAG-----	879
Db	910	TGGGAAGAGAGAAAGATCTTAAACAGAGAAACAAATTAAGTACCTTGCGGCTT	969
OY	880	-----CCGGCGGGAAGATTTGTTGGGATTAATGTTACGGCAAAAGATGGAACGTTCA	935
Db	970	ATTACTGAGATCTAAACAGAGCAATCAAAATGAAATGCAAGCTGGGATTTCAACAGA	1029
OY	936	GGAACATTTGAGAGAGTGTAAACGTTTTTCTTCGCGCGGAAACAGACATTTCTAATCT	995
Db	1030	AGATGTGATTTAGGATGATGAAGTTATTTTACTTTGCAAGTATGAGAACATCATAGTCTT	1089
OY	996	GCTGACGTGAGACACATCTTGCTATCCATGACACCCGGAAGTGGACGCGCAAGACAGGA	1055
Db	1090	GCTTACTTGGACATTTATTTGTGCTAAGCATGACCCAAATGGCAGAGAGACAGAGAGA	1149
OY	1056	TGAGATCTCAGAGTCTGCGGCTCACGTAATGTCCTTACCAAGACATGTCGTTAAGCT	1115
Db	1150	AGAGGTTTGAACCACTTTGG--AAGAACCAACACAAATTAAGTATGATCTTGAACCGCTT	1206
OY	1116	TAAACGTTGAGATGATCTTGAAGAGCTTTTAGGTGATTCACCAATATAGTAGTAC	1175
Db	1207	CAGACTGTAAACCATGATTTCTACATGAGTCTTATGTTGTATCCACTGGCAACCTTCT	1266
OY	1176	GATTCGACGCGCTAAATCGATGTGAAGCTTAGAGGGTACAAATATCCATGTGCAACGGA	1235
Db	1267	AACCAAGAAACCTTTAAGGAATGGAAGCTCGGTGAATCAAAATATCTTCGACGAGTGTGA	1326
OY	1236	GCTTCTAAATCCCAATCATAGCGGTGCATATACCAAGCAATTTGGGCTTAATGACGTGA	1295
Db	1327	GCTCTCTTCTGCGCGTATCTTCAATTCACATATATCCGACATTTTGGGAAAAAGACGAAAG	1386

```

QY 1296 CGAATTCATTCAGAGCTGCTTTGGATGAGAGTGCAGCGGTCGCCAAACACCCCGTTGG 1385
Db 1387 CGAATTCATTCAGAGAGTTTGGCAAGGCGATCTCCAGGCAACCAAGGATCAGGCTGC 1446
QY 1356 CTTATATACGCTTTGGGCTCGGAGTTCTGTACATGATTTGGTCAGAAATCTTGCTATCTTCA 1415
Db 1447 TTTCTTTCGGTGGAGGGGGGCCACAGATCTGCATCGGCGAGAGCTTTGCGTTGCTGGA 1506
QY 1416 GGCCAATTTAGCTGCTGCTTTAATGATGCACAGCTTACCTTCACTTGCTGCTCTTACTTA 1475
Db 1507 AGCCAAATGAGCGCTTATGACCAATCTCCAGCGCTTCTGCTTGAAGCTCTACCATCTTA 1566
QY 1476 TCAGCATGCACTTACCGTCTTATGTTGCTTTATCTCTGAAATGATGTGACCAATCAGCTT 1535
Db 1567 CACCCAGCGCGCTTACCGTGAATGATGACATGCAACCTTCAGACGAGTCTCAGATTAAGCT 1626
QY 1536 C 1536
Db 1627 C 1627

RESULT 9
US-10-424-599-85977
; Sequence 85977, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: LA ROSA Thomas J
; APPLICANT: KOVALLIC David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 85977
; LENGTH: 1922
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48650C.1
US-10-424-599-85977

Query Match 16.3%; Score 254.2; DB 13; Length 1922;
Best Local Similarity 50.6%; Pred. No. 1,2e-67;
Matches 759; Conservative 0; Mismatches 718; Indels 24; Gaps 5;

QY 72 AGTGAAGGTATGCTCTGTTATGATGTGAGACCAAGAAAGATTGAGACATTTCTCTTA 131
Db 107 ACTGAAGATGCTGAAGTCTGCTGTGCTTATGAGCCAAAGGTTGAAAGCTTCTTAAGC 166
QY 132 ACAAGGAATTCGAGGTCTCTCTTATCATTTCTTCAATGGAATGTTAAGAATCTGTGG 191
Db 167 GCAAGGTTTATGATGTGACCAATCTCACTTCTCATACCAACTCCAAACCAACCTTTT 226
QY 192 AATGATGTTAAGGTTCTTCTCATCTCATGCTTCTCTCTCAATATCTTCTCTAAGT 251
Db 227 GCAACAAACACGCTCAATTCCTCAACCTTTTCTCTATCCATGATGTGCTCTGAT 286
QY 252 TCTCTTTTTCACATCACTGAGAAAATCTACGCTGTCATTTCTGTTGGTGG 311
Db 287 ATCTCTCTGTTTACATCACTATGACAAATATGACAAAGAAATCTTCTTTTGGGAAG 346
QY 312 TCCAACTTCCGTTAAGGTAGCGATGCTGATTTGATTCAGAGATCTTCTTAAGTC 371
Db 347 TGAACACCAAGATGATCATTTAGCGAACCAAAATCACTCAAAAGATCTTCAACAACAT 406
QY 372 TGAAGTTTACGAGAAATGAAGCTCACCTTTGGTTAAACAACATGAAGGATGAGAT 431
Db 407 TCATGATCTCCGAAAGCCAAAGTTTAGTGAATATGCAATGTTCTTGT---GCTGGCT 463
QY 432 ACTTATGCTCAAGGTGAAAATAGGCTCATATCATGAAAAATCATTAAGCCCTATTCA 491

```

Db 464 GTTAATATAGAGGTGCAAGTGGGCTAAACATCGAAGAAATTATGACCACGATTTTC 523
Qy 492 TATGAGAACTTTAAGTCTTGTACCAAGTTGTGTGAAGAGTGTACTGATATGTTGGA 551
Db 524 CTCAGAAAAATGAAAAATATGCTACCGCATTTTTCACAAAGCTGCCATATATGATTAG 583
Qy 552 TAAATGTCCTGATATGATATGAGAAAACGTGAGTGTAGATATGATATGATGATG 611
Db 584 CATGTGAAAGGAAAGTGTGTATGATGAGTGAAGAAATGTGATGACATTTGGGCTTCT 643
Qy 612 TCAGATTTGACTGAGATGTTATTTAGTGAACAGCTTTTGAAGTATGATGAGATGAG 671
Db 644 TCAGATTTTAACTCGTATGTTATTTTCACAAAGCATTTGGAAGCAGTATGCAAGAG 703
Qy 672 TCGAGCAGTTTTCGACTCAAGCTCAACAAAGCTTTTGTGCTGAGCTTTTGAAG- 730
Db 704 AGAAAAATTTTTCGAAAATCGAGATGCAAGGCTATCTTTTAAATGCGAGAAAGTACA 763
Qy 731 -----AAGCTTCAATTCCTGCTATGATTTTTCGACAAAGGGAATTTGAAGT 781
Db 764 GAACATACCAATATTTGCGGCATCTACGTACAACTACCAAGAGATGAGACAAATTGA 823
Qy 782 CTCGGAATTTGACAGAGATGAGAAAGTGTGTTGAGAGCTGATGAGCGCGGAGA- 840
Db 824 AAGAGATACGAGATTCATTTGAGGGTATCAAAAAAGAGAGAAAGCCATGAGAA 883
Qy 841 -----CAAAACCTATAGATGAGAAAGGGAAGAAATGTAAGAGCGCGCGGAGATTT 896
Db 884 TCGTGAACCTTATGAGAAATTTATTAAGCATCTTTGGAATCAATACAGAGAAAT 943
Qy 897 GTTGGGATT--AATGATTCAAGGCAAGAAATGTGACGGTTCAGACATTTGAGAGATG 953
Db 944 CCGAGGACATGGAACAGTAGGGCTGTGGAATGACCAAGCAAAAGTATTTAGGATG 1003
Qy 954 TAAAGCTTTTCTTGGCGGGGAAACAGCAACTTCTATCTGTCAGCGTGAAGCAT 1013
Db 1004 CAGCTATTTTACCTGCGAGGCAAGACACCTTCTCTGCTGTTGGACATAGT 1063
Qy 1014 CTGCTATTCATGACCCCGAGTGGCAGGCAAAAGCAGTATGAGGTCTCAGGGCTG 1073
Db 1064 CTATATGGCTAGGTATCCAGATGCGACAGCAAGAGGAGCAAGATTTTTCAGTTT 1123
Qy 1074 CGGCTCAGGTATGTCCTTACCAAGACATGTGTTAAGCTTAAAGCTTGAAGTAT 1133
Db 1124 TGGAAACCAAAAT--CCAAACATTTGAGGTTAAGTGAAGCTTAAATGTGACATGAT 1180
Qy 1134 CTTGAACGATCTTTAAGTGTATCCACCAATGATGCTACGATTCAGCGCGTAAATC 1193
Db 1181 TTTAATACAGGTTCTCAGGCTATATCTCCAAACAATTTCTTATGCGCGCACTCAAA 1240
Qy 1194 GGAATGAGAGTGAAGGCTACAAATCCCATGTGGCAGCGAGTCTTATCCCAATCAT 1253
Db 1241 GGAATGAACTTGAATACTATCTACCTGCGAGAAATACGATTTCCATGCGCAATAT 1300
Qy 1254 AGCGTCCATCATGACCAAGCATTGGGGTAAATGACGTGAACGAATTCATCCAGCTCG 1313
Db 1301 ATTATTCATCATGATGATGATATCTGGGGTGAATGACAAAGAGTTCAAACTGAAAG 1360
Qy 1314 GTTTGGGATGAGAGTGGCGGCTGCTGCCAAACCCCGTGGCTTCAATCCGTTGGCCT 1373
Db 1361 GTTTTCTGAAGAAATTTGCTAAGGCAACAAAGGCCAAATTTGTTTATCATTTGGATG 1420
Qy 1374 CGAGTTGATCATGATGATGATGATCTGTATATCTCAGGCAAAATTTGACACTCG 1433
Db 1421 GGGTCTGAAATATGATGATTTGGCAAACTTTGCTTAAATGGAACAAAGATGATATC 1480
Qy 1434 TGTAAATGATCAACGCTTCACTTCACTTGGCTTCACTTATAGATGACATCTTACCT 1493
Db 1481 ATTGCTTGGCAACCTTCTCATTTGAGCTTTCGCGGTATATGATGATGATGATGATG 1540
Qy 1494 CCTATGTTGCTTATCTCAACATGATGACCAATCACTTCCGAGGATTTGACCAATCA 1553

Db 1541 TGTGCTTAGTTTGACGCCAAACGTGGGCGACATGCTTTTGCATATAATGATATGA 1600
Qy 1554 T 1554
Db 1601 T 1601

RESULT 10
US-10-424-599-126502
; Sequence 126502, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 126502
; LENGTH: 1872
; TYPE: DNA
; ORGATISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_85238C.1
US-10-424-599-126502

Query Match 16.0%; Score 250.8; DB 13; Length 1872;
Beet Local Similarity 50.6%; Pred. No. 14e-66;
Matches 749; Conservative 0; Mismatches 707; Indels 24; Gaps 5;

Qy 90 GTTATGTGAGACCAAGAAAGATGGAAGAACTTCTTAAACAGAAATTCGAGTCC 149
Db 144 GGTGTGGGTGAGCCCAAGAGATGAGAGCGCTCAAGACAGGGTATCCAGAGAA 223
Qy 150 TCTTATCATCTTCTTCAATCGGAAATGTTAAGAACTTTGGAATGATGTTAAAGCTTC 209
Db 224 TTCTTACCGCTTGTATGATGAGCATCAGAGATGATGATGATGATGATGATGATGATG 283
Qy 210 TTTTCATCTATG--CCTTCTCTCAAAATTTCTCTAGAGTCTCTCTTTTAAACA 266
Db 284 ATCCAAACCTATGATCTCAGCTATATGACATGACACCCGATGTGCTTACCTGT 343
Qy 287 TCACTGAGAAAAATCTACGCTGCTACATTTCTGTTGGTTCGGTCCAACTTCCGGTT 326
Db 344 TCAACCACTGCTAAATACGCTAAGATGATGATGATGATGATGATGATGATGATGATG 403
Qy 327 AACGTTACCGATCTGATTTGATCAGAGATCTTCTTAACTGAGTCTTACAGAA 386
Db 404 ATTCACTGATCTGACAAATTCAGAAAGATGCTAATGATGATGATGATGATGATGATG 463
Qy 387 GAATGAGCTACCTCTTGTGTTAAACAATTGAAAGGAGTACTGTTAGTCTCAAAG 446
Db 464 GCCCGACATGATGACATTTTCA--AGCTTCTGACATCAGGGTTGGCAATTAATGATG 520
Qy 447 TGAATAATGGGCTCATCATGAAATAATCATTAAGCTTCTTCAATGAGATCTTAA 506
Db 521 CGACAGTGGGCTTAAACAGAAAGATCGAGATCCAGATTCATGATGATGATGATGATG 580
Qy 507 GTTGTGTATCAGATGTTGTTGAAGAGTGTGATGATGATGATGATGATGATGATGATG 566
Db 581 GCTCTGTATCCGATATTTTCCAGAGTGTGATGATGATGATGATGATGATGATGATGATG 640
Qy 567 GTTATCGAA--AACGTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 623
Db 641 GTTATCTTCTGCTTAAATGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 700
Qy 624 TGAAGATTTATATGATGAAACAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 683
Db 701 AAGGACGTTCTCTCGTGAAGCTTTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 760

QY 684 TCAGCTTCAAGCTCAAAAGTCTTTGTCGTAGCTTTTCAAAAGTCTTCAATCC 743
 Db 761 CGAATTTCAAGGAAATGATTCAGCTTCAATGACGCTTTAAGTTGCTTCAATCC 820
 QY 744 TGGCTATAGATTTTTCGACAAAGGAAATTTGAAGTCTCGAAAGTTAGCAAGAGAT 803
 Db 821 AGGTACAGTTTCTCCAAAGCATTCACAAAGAGATGAAAGCAATTTGCAAGAAAT 880
 QY 804 AAGGAAGTGTCTTGAAGCTGATAGAGCGGCGGAGAAAGCAAGCTATAGATGAGAG 863
 Db 881 ACAGAGATCATTTATGTATCATCAACGAAAGATTTAAAGCAATCAAGAGAGAGAGCC 940
 QY 864 GGAAGATGTAAGAGC-----CGCGCGGAGAGATTTGTTGGATTAATGAT 911
 Db 941 TACTTACATGACTTTTGAAGCACTCTTGAATCAATTAACAAGAAATCTGAAAAAG 1000
 QY 912 TCAGGCAAGAAATGTACGCTTCAAGCAATTTGTGAGAGAGTGAAGCTTTTCTTCC 971
 Db 1001 TAGGTGTGAGAAATGATTAAGGAAAGTGAAGAGTGAAGAGTAAATTTTCTTGGC 1060
 QY 972 CGGAAACAGCAACTTCTAATCTGCTGACGTCAGCAACATCTTGTATCCATGCAACC 1031
 Db 1061 AGGCGAGAGAAATGCAAGAAATGCTGCTGACATTTGTTATTAACAGCAATCC 1120
 QY 1032 GAGTGGCAGAGCCAAAGCAGTGAAGAGTCTCTAGAGTCTGCGGCTCAGCTATGCC 1091
 Db 1121 AGATGGCAAGAAAGCTGCGGAGAGAGTCTTCCAGTGTCCG---GAATGAAAGCC 1177
 QY 1092 TACCAAGACATGCTTAAGCTTAAAGCTTGAATGATCTTGAACAGCTTTAAG 1151
 Db 1178 AGATTAATGAAGATGTCTAAGCTTAAAGCTTGAATGATCTTGAACAGCTTTAAG 1237
 QY 1152 GTTGTATCCACCAATGATGATCAATGATGACGCTTAATGATGATGATGATGATG 1211
 Db 1238 ATTAATCCACCAAGTGTATGATGATGATGATGATGATGATGATGATGATGATG 1297
 QY 1212 GTTCAAAATCCCATGTCGACAGAGATTTCTAATCCCAATCAATGAGGCTCAATGACCA 1271
 Db 1298 GCTTACAAATCCCGCAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1357
 QY 1272 AGCCATTTGGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1331
 Db 1358 GAGATTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1417
 QY 1332 GCGTGTGCAAAACCCCGTGGCTTCAATCCGTTTGGCTTCGAGATTCGATGATG 1391
 Db 1418 AAGGCAACAAAGGAGAGCTTCTCACTTGGCATTTGATGAGGCTCTCGACTCTGAT 1477
 QY 1392 TGGTCAAGATCTTGTATTAATCTCAAGGCAATGACATGCTGATTAATGATGACAGCTT 1451
 Db 1478 AGGCAAACTTTGGCTTGTAAAGCAAAAGATGATGATGATGATGATGATGATGATG 1537
 QY 1452 CACCTTCACTTGGCTTCTAATTAATGATGATGATGATGATGATGATGATGATGATG 1511
 Db 1538 CTCCTTCATTTTCCCTCTCTATGCTATGCTATGCTCTCTCTCTCTCTCTCTCTCT 1597
 QY 1512 TCAAGATGTGCAACATCACTTCGGAATGACCAAT 1551
 Db 1598 TGAAGGTGGGCTCATCTCTTTTAAAGCAATCTTGAAT 1637

RESULT 11
 US-09-938-842A-92
 ; Sequence 92, Application US/0993842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Krepis, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE

FILE REFERENCE: SRRP1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938, 842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 92
 ; LENGTH: 1557
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-92

Query Match 15.4%; Score 241.2; DB 9; Length 1557;
 Best Local Similarity 49.2%; Pred. No. 1.2e-63;
 Matches 754; Conservative 0; Mismatches 768; Indels 12; Gaps 4;

QY 10 GAAAGTACAGCTGCTTCAATTCGAAAGTCTTGTCTGTCTGATTAATCTTAAGCTTGTG 69
 Db 4 GAACTTAATAGCAATCAATCTTAAATCTTCTTCACTTCTGTTGTTCCAG 63
 QY 70 ATAGTGAAGGATGCTCTGTTAATGCTGAGACCAAGAAAGATTAAGAACTTCTCT 129
 Db 64 ATATGAAAGCTTGTGATCTCTTCTTGTGCTCAATGATGATGATGATGATGATGATG 123
 QY 130 AAACAAGAAATTCAGAGTCTCTTATCATTTCTTCAATGGAATGTTAAAGCTTGT 189
 Db 124 AAACAAGAAATTCAGAGCAACAAAGTCTTGAACGAACTCAGAGATTAAG 183
 QY 190 GGAATGCTTAAAGCT---TCTTCAATGCTTAAAGCTTCTCTCAATATCTTCT 246
 Db 184 AAGATGAAGAAAGAGCTGACCTTGTGTCTGATCCCAATGATTAATTTTCTCT 243
 QY 247 AAGATGCTCTCTTCTTCAATCATCTGAGAAATTAAGCTGCTTCAATTTCTGTTGG 306
 Db 244 CGTGTCTTCTCAATCATCAATGATGATGATGATGATGATGATGATGATGATGATG 303
 QY 307 TTGCTTCACTTCTGCTTAAAGCTGATGATGATGATGATGATGATGATGATGATGATG 365
 Db 304 ACTGAAACAAACCAATATATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 363
 QY 366 --TAACTGATGCTTCAAGAGAAAGTAAAGCTCCTTGTGTTAAACAACTTGAAGC 423
 Db 364 AGCAAGTGTGTTTCACTATTAATACGCTGAAGAGAGCTGAGCTTCAATCTTCTGCT 423
 QY 424 GATGATCACTTATGCTCAAGGATTAAGGCTGATCAATGAAATTAATGATGATGATG 483
 Db 424 AAAGACTTCTCTTATCCAAAGTATGATGATGATGATGATGATGATGATGATGATG 483
 QY 484 ACTTTCATATGAGAAATCTTAATGATGATGATGATGATGATGATGATGATGATGATG 543
 Db 484 GCTTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 543
 QY 544 ATGATGATTAATGCTCGATTAATGATGATGATGATGATGATGATGATGATGATGATG 600
 Db 544 ATATTTGAAGATGAGAAACAGAGAAAGTGTGAAGTGTGATCAATGATGATGATG 603
 QY 601 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 Db 604 AGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
 QY 661 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
 Db 664 TATGCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 723
 QY 721 GCTTTCAAAAGTCTCAATCTGCTTAAAGATTTTTCGACAGAGAGGAAATTTGAG 780
 Db 724 TCTCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 783
 QY 781 TCTGCAAGTGAAGAGATTAAGAGATGATGATGATGATGATGATGATGATGATGATG 837

Page 11

```

SEQ ID NO 92
LENGTH: 1557
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-92

Query Match
Best Local Similarity 49.2%; Pred. No. 1.2e-63;
Matches 754; Conservative 0; Mismatches 768; Indels 12; Gaps 4

15.4%; Score 241.2; DB 11; Length 1557;

10 GAAAGTAGCAGCTGCTTCAATTCCAAAGGCTCTTGTCTGTCTGAATCTTAACTCTGTA 69
Db GAACTTAATAGCAGCAATCAATCTCTTAACATCGCTCTTCTACTCTTCTGTTTCCAA 63
QY ATAGTGAAGGATGATCTCTGTATATGAGGAGACCAAGAAAGATTGAGACATTTCTCT 129
Db ATATGGAAGCTTTGATCTCTTGTGATCTCTTGTGGTCCATTGATCTATCAAAAAGATTCA 123
QY AAAACAAGATTCGAGCTCTCTTATCATTTCTTCATCGAAATGTAAAGACTTTGT 189
Db AAAACAAGATTCGAGCAACCAAGTCTGTATCGAAACCTCACCGAGTAAG 183
QY GGAATATCTTAAAGCT---TCTTCATCCATATGCTTTCTCTCAATATCTTCTCT 246
Db GGAATATCTTAAAGCT---TCTTCATCCATATGCTTTCTCTCAATATCTTCTCT 243
QY AAGATTAAGAAAGAGCTGACCTTGTGTGTGATCCAAATTCCAATGATATTTCCCT 243
Db AAGATTAAGAAAGAGCTGACCTTGTGTGTGATCCAAATTCCAATGATATTTCCCT 243
QY AGAGTCTCTCTTTTACCATCACTGAGAAATACTACGCTGTACATTTCTGTTGG 306
Db CGGTCTTCCAGATACCAATGATATGATCTCAATACGAGATACATTTCTATCTGG 303
QY TTGCTCAACTTTCCGGTTACAGGTACGCTCTCTGATTTGACAGAGATCTTCTC- 365
Db TTGCTCAACTTTCCGGTTACAGGTACGCTCTCTGATTTGACAGAGATCTTCTC- 363
QY ACTGGAACCAACCACTATATACATCTCAATATAGAGCTGACGGAACAGTCTTGTG 363
Db ACTGGAACCAACCACTATATACATCTCAATATAGAGCTGACGGAACAGTCTTGTG 363
QY --TAACTGAGTTCTACGAGAGATGAAAGCTCACCTTGTGTTAAACAATTGAAGC 423
Db AGCAATTCGCTTACATTAATACCGGTGAAGAGCTGAGGCTTCAATCTTTCGGT 423
QY GATGACATACTAGTCTCAAAGGTGAAAATGCGCTCATCATGAAAAATCATTAAGCC 483
Db AAAGGACTTTCCTTATCCAAAGTATGATTTGATGCGCATACGAACTTGAACCT 483
QY ACTTTCAATGAGAGATCTTAAGTGTCTTGTACAGTGTGTGTAAGAGTGTACAT 543
Db GCTTTCTCATGATGATGCGCTCAAGGCCATGACACACCGAATGGAGATGACCTTA 543
QY ATGATGATTAATGTCGATTAATTATCA--GAAAACGCTGAAGTTGAGTATGTC 600
Db ATATTTGAGAGAGTGGAAGAAACAGAGAAAGATGATGAGTTGATCAAGATTGAGATA 603
QY TATGATGCTTTCAGATTTTGACTGAAGATGTTATAGTAAACAGCTTTTGGAGTAC 660
Db AGCAAGAGTTCATTAATGACCGCTGACATTAATAGCACTAATCGGTTTGAAGT 663
QY TATGAAGATGTCGACAGCTTTTTCGACTCAAGCTCAACAAATGCTTTTGTGCTGAA 720
Db TATGCGGAAGGATCGAATTTGTGTAGTCAACAAACAGAGCTTGAAGATTATATTAAGT 723
QY GCTTTCAAAAAGTCTCAATCTGCTGATTAAGATTTTTCGACAAAGGGAATTTGAG 780
Db TCTCTCACTAAGCTTTTCAATCCCGAATCAATACCTTCTTAAGCTTACCAACTTAA 783
QY TCTCGAAGTTAGACAGAGAGATTAAGAGTGTGTTGAAGCTGATAGA--GCGCGG 837
Db CTATGGAGAGCTCATTAAGAAAGTGAAGAACTGATCAAAAGATCATATGATTTCAAGCTA 843
QY AGACAAACGCTATATATGAGAGAGGAGAAATTAAGAGCGCGGCGGAGAGATTG 897
Db AAATCAAAATGTAAGACTTATGAGCTATGAGGAGCATTTCTAGGGGTATGTTACTGCT 903
QY TTGGGATTAATGATTCAGGCAAGATGTGAAGGTTTCAGCATTTGAGAGAGTGTAA 957
Db GCGAAATCTAAGAGTATGAGAGAAAGATGAGAAATGATGATGATGATGAGAAATGCAAG 963

```


Mon May 24 08:10:08 2004

us-09-992-901-1.rnpb

Page 14

ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3608-017-E5_FLI
US-10-425-114-24276

Query Match 14.7% Score 230.2; DB 13; Length 1808;
Best Local Similarity 50.2%; Pred. No. 3.6e-60;
Matches 641; Conservative 0; Mismatches 613; Indels 23; Gaps 2;

```
QY 284 ACGGTCATCATTTCTGTTGGTTCACATTTCCGTTACGTTACCGATCTCTG 343
DB 343 ACGGCAAGACGTCTTCACTGGTTCGCGCCCGTCCCAAGGTGACCTCGACCG 402
QY 344 ATTGATCAGAGATCTTCTGAAGTGAAGTTTACGAGAGATGAAGCTCACCTT 403
DB 403 ACTGGCGAGGAGCTGATGCGCACAAAGTTGCGCACTTGGAGAGCCCAAGTTCCCG 462
QY 404 TGGTTAAACAACCTTGAAGCGATGGAATTAAGTCTCAAGGTGAAAATGGGCTCATC 463
DB 463 CGCTCAACCAAGCTGTTCTCCGACGCTGTCGCCAACCAAGGCGAGAGTGGGTCAAGC 522
QY 464 ATGGAATAATCATTAACCTTACTTCTTATATGAGAAATTAAGTTGCTTACCAATG 523
DB 523 ACGAGAGATCCCTCAACCTTGGTTCACCTGAGAGCTCAAGCTCATGCTGCGGCGT 582
QY 524 TGTGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 583
DB 583 TCTCTGCTGCTGCTGATGAAGAGCTGTCAGCGCATGGGCGAGTGGCTTGGCCCGACG 642
QY 584 AAGTTAGTAGTAGTGTCTATGAGTGTTCAGATTTTGAATTAAGTAAATGTTAGTAA 643
DB 643 CGTGGAGCTGAGCGTGAACCCGAGACTTCAACCTCAACCGGAGTGTCACTTCTCGCA 702
QY 644 CAGTTTGGAGTAGTATGAGTGAAGATGTCGAGCAATTTTTCAGCTTCAAGCTCAACAA 703
DB 703 CCGGTTCCGACAGCTACCTACCTGAGAGAGAAAGTTTTCAGCTTCCAGGCGAGCAAG 762
QY 704 TGGCTTTTGTGCTGAAGCTTTTCAAAAAGTCTTCACTTCTGCTATGATTTTTCGA 763
DB 763 CTGAGCGCTCATGCTCATCATGATGACAAATTCGCCCTTCCGATATCATGCTTGCCTA 822
QY 764 CAAAGAGGAATTGAAGTTCGGAATGGAACAAGATTGAACAAGATGAAGTGTGTTGAAGC 823
DB 823 CGAATAACAAACCAAGATGCGCCAAATCAAAAGCAGATCGATCGATCTTGGGGCC 882
QY 824 TGATAGAGCGCGAGACAAACGCTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 870
DB 883 TGATCGGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 942
QY 871 -----TGTAAGAGCGGCGGCGAGAGATTGTGGATTAATGATTCAGGCAAG 921
DB 943 GCTTACTGCTGAGTCAACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1002
QY 922 AATGTGACGTTTCAGAGCATTTGAGAGAGTGAAGCTTTTCTTCCCGGAGAAACAG 981
DB 1003 GAT-TGAGATGAGAGAGATGATGAGAGAGTGAAGCTGTTCTTACTTCCGCGAATGAG 1061
QY 982 ACAACTTCTAATCTGCTGAGTGAAGACATCTTGTATCCATGACCCGAGAGTGGCAG 1041
DB 1062 ACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1121
QY 1042 GCCAAGACGCTGATGAGAGTCTCAGAGGTCTGCGGCTCAGTATGTCTTACCAAGAGC 1101
DB 1122 GACCGTGAAGAGAGAGAGTCTCGACTGTTGGGAGAGAGAGAGAGAGAGAGAGAGAG 1181
QY 1102 CATGCTTAAAGCTTAAACGTTGATGATGATGATGATGATGATGATGATGATGATG 1161
DB 1182 GAGCGTGAAGTCCCTCAAAACGATGACATATCTTACAGAGTCTTCCGCTGTACCCG 1241
QY 1162 CCAATAGTAGCTAGATGACAGGCTAAATCGATGATGAGAGTGAAGGTAACAAATC 1221
DB 1242 CCGGCAATCGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1301
```

```
QY 1222 CCATGTGACAGAGAGCTTCTATCCATCAATGACGCTGCATGATGACCAAGCCATTGG 1281
DB 1302 CTTGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1361
QY 1282 GGTATGACGTGAAGAGATTTCAATCCAGCTGCTTGGAGAGAGAGAGAGAGAGAGAGAG 1341
DB 1362 GGAAGCAGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1421
QY 1342 AAACACCCGTTGCTTATACGCTTGTGCTTGGCTTGGAGTGTGATGATGATGATGATG 1401
DB 1422 AAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1481
QY 1402 CTGCTATATCTTACAGGCAATTTGACACTGCTGATATGATGATGATGATGATGATG 1461
DB 1482 TTGCGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1541
QY 1462 TTGGCTCTATCTTATCAGATGACACTACGCTCTTATGTTGCTTATCTTCAACATGT 1521
DB 1542 CTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1601
QY 1522 GCAACATACCTTCCG 1538
DB 1602 GCCAGATCAAGCTCAG 1618
```

Search completed: May 22, 2004, 22:05:02
Job time : 743 sec

CC increases resistance to insects in plants. The present sequence
CC represents the Arabidopsis bael protein sequence, which is used in the
CC exemplification of the present invention
XX
SQ Sequence 520 AA;

Query Match 100.0%; Score 2730; DB 3; Length 520;
Best Local Similarity 100.0%; Pred. No. 3.3e-258; Indels 0; Gaps 0;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEESSSWFIPKVLVSLVSLVYKMSLMMRPRIEHSKQIGRPYHFIQNVK 60
DB 1 MEESSSWFIPKVLVSLVSLVYKMSLMMRPRIEHSKQIGRPYHFIQNVK 60
QY 61 ELVGMMLKASHPMFPSHNLPRVLSFYHWRKTYGATFLVWGPTRLTVAADPDIREI 120
DB 61 ELVGMMLKASHPMFPSHNLPRVLSFYHWRKTYGATFLVWGPTRLTVAADPDIREI 120
QY 121 FSKSEFEKNEAHPVYKQLEGGDLISLGEKMAHRRKIISPTFMENLKLVVYVLSVT 180
DB 121 FSKSEFEKNEAHPVYKQLEGGDLISLGEKMAHRRKIISPTFMENLKLVVYVLSVT 180
QY 181 DWVDMKSDKLSNGEVEVDVYEMFQILTEDVISRTAFGSSYEDGRAVFRLOAOQMLCAE 240
DB 181 DWVDMKSDKLSNGEVEVDVYEMFQILTEDVISRTAFGSSYEDGRAVFRLOAOQMLCAE 240
QY 241 APOKVTIPGYRFPFRGNLKSRLDKETIRKSLKLIERRRONAIDGSECEKPEAKDL 300
DB 241 APOKVTIPGYRFPFRGNLKSRLDKETIRKSLKLIERRRONAIDGSECEKPEAKDL 300
QY 241 APOKVTIPGYRFPFRGNLKSRLDKETIRKSLKLIERRRONAIDGSECEKPEAKDL 300
DB 241 APOKVTIPGYRFPFRGNLKSRLDKETIRKSLKLIERRRONAIDGSECEKPEAKDL 300
QY 301 GLMIQAKNTVODIYECKSPFFAGKQTTNNLITWTTILSMPEMOAKADEVLAVCGS 360
DB 301 GLMIQAKNTVODIYECKSPFFAGKQTTNNLITWTTILSMPEMOAKADEVLAVCGS 360
QY 361 RDVPTQDVAHVYKLTSMILNESLRYPVATIRAKSDVYLGKTKPCGTETLIPITAV 420
DB 361 RDVPTQDVAHVYKLTSMILNESLRYPVATIRAKSDVYLGKTKPCGTETLIPITAV 420
QY 421 HHQDAIWDVNEFNPAREADGVPRAPAKHPVGFIFGLGVETCIQONLAILQAKLTAVM 480
DB 421 HHQDAIWDVNEFNPAREADGVPRAPAKHPVGFIFGLGVETCIQONLAILQAKLTAVM 480
QY 481 IQRFTHLAPTYOHAFTVLMILYPOHGAFTFRRLTNED 520
DB 481 IQRFTHLAPTYOHAFTVLMILYPOHGAFTFRRLTNED 520

RESULT 2

AB08079 standard; protein; 527 AA.

AC ABB08079;
XX 10-SEP-2002 (first entry)
DT 10-SEP-2002 (first entry)
DE Maize cytochrome P450, CYP72A1.
KM Maize; cytochrome P450, CYP72A1; CYP92A1; plant; gene therapy; anticodote.
XX Zea mays.
OS
XX Key Location/Qualifiers
FH 200..527
FT /note="this fragment is not provided in the protein
sequence indicated in the sequence listing"

XX US6380465-B1.
XX 30-APR-2002.
XX 11-JUL-1999; 99US-00351229.
XX 12-JUL-1998; 98US-0092596P.

XX (KENT) UNIV KENTUCKY RES FOUND.
PA Barrett M;
XX WPI; 2002-470227/50.
DR N-PEDB; ABL60758.
XX Nucleic acid molecules encoding a Zea mays cytochrome P450 CYP72A1,
PT useful for the production of plants resistant to heterocyclic,
PT sulfonylurea and substituted urea herbicides and organophosphate
XX insecticides e.g. Classic and Pursult.
PS Claim 13; Fig 1; 21pp; English.

CC The invention relates to isolated nucleic acid molecules encoding Zea
CC mays (maize) cytochrome P450 designated CYP72A1 and CYP92A1. The
CC polypeptides can be expressed by standard recombinant methodology. The
CC nucleic acids may be used in genetic engineering protocols to transform
CC plants and other eukaryotes e.g. yeast, maize (especially), soybean,
CC beet, tobacco, wheat, barley, poppy, rape, sunflower, alfalfa, sorghum,
CC rose, carnation, gerbera, carrot, tomato, lettuce, chichory, pepper, melon
CC and cabbage. They may be used in this way to confer protection against
CC heterocyclic herbicides, sulfonylurea herbicides, substituted urea
CC herbicides and/or organophosphate insecticides. The present sequence
CC represents the Z. mays CYP72A1

SQ Sequence 527 AA;

Query Match 42.6%; Score 1163.5; DB 5; Length 527;
Best Local Similarity 42.5%; Pred. No. 1.1e-104; Indels 19; Gaps 5;
Matches 224; Conservative 103; Mismatches 181; Indels 19; Gaps 5;

QY 1 MEESSSWFIPKVLVSLVSLVYKMSLMMRPRIEHSKQIGRPYHFIQNVK 60
DB 8 MLREASPMALAGA-VASVSLMLVAWTLBAMWMTPMWLDALNAQINGRRLFTGDDLR 66
QY 61 ELVGMMLKASHPMFPS-HNLPVLSFYHWRKTYGATFLVWGPTRLTVAADPDIREI 119
DB 67 ETARVNRARAKKPLDGGHDTTPRAYOPMNSTIYEGKLSFTFMGPTRPVIDPBLVKE 126
QY 120 IFSKSEFEKNEAHPVYKQLEGGDLISLGEKMAHRRKIISPTFMENLKLVVYVLSVT 179
DB 127 VLSNFGHFGKPRSSRIGRLANGLVNHDDGKMAKRRRIINPAVHHRKIGAMPVESGCC 186
QY 180 TMVDMKSDKLSNGEVEVDVYEMFQILTEDVISRTAFGSSYEDGRAVFRLOAOQMLCA 239
DB 187 IEMITRMDSMSKSSSEIDVWPEQNTGVDVSRITAFGSSYEDGRAVFRLOAOQMLCA 246
QY 240 EAFQKVTIPGYRFPFRGNLKSRLDKETIRKSLKLIERRRONAIDGSECEKPEAKDL 299
DB 247 QSVQITIFIPGVWFLPTNNRRRAIDVIRIKLIEIKGEKDT-----KNEITNKDGL 300
QY 300 LGMLTQAK-----NTVODIYECKSPFFAGKQTTNNLITWTTILSMPEMOAK 349
DB 301 LGMLTQAK-----NTVODIYECKSPFFAGKQTTNNLITWTTILSMPEMOAK 360
QY 350 ARDEVYLRVCGSRDVPYKQLEGGDLISLGEKMAHRRKIISPTFMENLKLVVYVLSVT 409
DB 361 ARDEVYLRVCGSRDVPYKQLEGGDLISLGEKMAHRRKIISPTFMENLKLVVYVLSVT 419
QY 410 GTETLIPITAVHHQDAIWDVNEFNPAREADGVPRAPAKHPVGFIFGLGVETCIQONL 469
DB 420 GTETLIPITAVHHQDAIWDVNEFNPAREADGVPRAPAKHPVGFIFGLGVETCIQONL 479
QY 470 IIOAKLTAVMIQRFTHLAPTYOHAFTVLMILYPOHGAFTFRRLT 516
DB 480 IIOAKLTAVMIQRFTHLAPTYOHAFTVLMILYPOHGAFTFRRLT 526

RESULT 3

AAU97096 standard; protein; 512 AA.

XX AC AAU97096;
 XX DT 24-SEP-2002 (first entry)
 XX DE Abscissic acid (ABA) hydroxylase CYP72A7.
 XX KM ABA hydroxylase; abscissic acid hydroxylase; transgenic; plant; CYP72A7;
 XX KM enzyme.
 XX OS Arabidopsis thaliana.
 XX PN WO200246377-A2.
 XX PD 13-JUN-2002.
 XX PF 06-DEC-2001; 2001WO-CA001756.
 XX PR 07-DEC-2000; 2000US-0251518P.
 XX PA (CANADA) NAT RES COUNCIL CANADA.
 XX PI Krochko JE, Cutler AJ, Abrams SR;
 XX PI WPI: 2002-519663/55.
 XX DR N-PSDB; ABKS0985.
 XX PT New isolated and purified DNA that encodes protein having abscissic acid
 XX PT (ABA) hydroxylase activity, useful for altering catabolism of abscissic
 XX PT acid in plants.
 XX PS Disclosure: Page 106-109; 117pp; English.
 XX CC The invention relates to an isolated and purified DNA (I) that encodes a
 XX CC protein having abscissic acid (ABA) hydroxylase activity. (I) is useful
 XX CC for producing a transgenic plant which involves introducing (I) into a
 XX CC genome of the plant or its part, and carrying out plant growth and
 XX CC development. (I) is useful for modifying catabolism of ABA or ABA
 XX CC analogues in a plant. ABA hydroxylase clone pB10-30-3 cDNA sequence is
 XX CC useful for identification of related sequences from other plant species.
 XX CC (I) is also useful for altering ABA and ABA analog catabolism in plants.
 XX CC The present sequence represents abscissic acid (ABA) hydroxylase CYP72A7
 XX CC

SO Sequence 512 AA;
 Query Match 40.2%; Score 1098; DB 5; Length 512;
 Best Local Similarity 40.2%; Pred. No. 2,8e-98;
 Matches 208; Conservative 117; Mismatches 166; Indels 26; Gaps 7;

QY 17 SYLIVIVKMSLW-NR-----PRKIEHESKQIGRGPYHPTGNVKELVOML 67
 DB 4 SYVALPLVAVVAVVWTRIVKMWIKPMLESSIKRQGLGTPTPLVGDIKENVMM 63
 QY 68 KASSHMPSPSHNLPRLVLSFYHMKIKYGAFLVWFGPTFRLLTVADPDLIREISKSEFY 127
 DB 64 EAKSKKINVTDDITPRLPLALMKLNSGKTFEITIGLPTTVINPEQIKENVKMYDF 123
 QY 128 EKNEAHPVLVKGEGDGLSLKGEKNAHHRKISPTFHENIKLIVPVVLSKSTDMVDKMS 187
 DB 124 EKASTPPLRLKLAG-GLASVKDDXNASHRRIINPAFHLEKIKMIPAFYHCCSHVCOM- 181
 QY 188 DKLSNGE--VEVDVVEWFQITLTDYISTAFGSSYEDGKAVFRLQAOQMLTCAAFOKV 245
 DB 182 EKLFTEKSPLEVDVWPMVMTADYISHTAGSSYKGGQRIPLQGLAELIIOAFKKS 241
 QY 246 FIVGYRFPPTGRLKSKLDEIRKSLKLIERBRONADIGEGBCEKPAKDLIGLMI- 304
 DB 242 YVIGSRFFPKSRKRAIDREVDVLRGIVSKRE-----KARAGSPANDDLLGILLE 295
 QY 305 -----QANKVTYQDIVEECKSFFFGKQTTNLLTWTITLLSMPEVQAKRDEVLAVC 358
 DB 296 SNSEBQSGNMGMSYEDVKCKLFFYFAGQETTSVLLVMTVLLSHHODQAPAREVMQVL 355

QY 359 GSRDVPFKDHVVKLKTLSMLNLSRLYPPIVATIRRAKSDYKLGKIPCGTELLIPII 418
 DB 356 GENKPKMESLNLKWTMTAFNEVNLPLFPVAGLKEVNVKEMKLGELTLPAGIQVLPFI 415
 QY 419 AVTHDQAIWGNVDVNEFNPAPFADGVPRAAKHVPVGFIPBGLSVRTICGQVLLLOAKTLTA 478
 DB 416 LVGRDTELMGDDADAFKPERFRDGLSKATQNVSPFPFGMGRICIGQNFAMLEAKMAMA 475
 QY 479 VTIORFTEHILAPYGOHAPTVLMLLYPOHGAPITFRRL 515
 DB 476 LITOKFSEFLSPSTVTHAPQVTWTRTFQFAHLIHLTL 512

RESULT 4
 ID AAU97097 standard; protein; 512 AA.
 XX AC AAU97097;
 XX DT 24-SEP-2002 (first entry)
 XX DE Abscissic acid (ABA) hydroxylase CYP72A14.
 XX KM ABA hydroxylase; abscissic acid hydroxylase; transgenic; plant; CYP72A14;
 XX KM enzyme.
 XX OS Arabidopsis thaliana.
 XX PN WO200246377-A2.
 XX PD 13-JUN-2002.
 XX PF 06-DEC-2001; 2001WO-CA001756.
 XX PR 07-DEC-2000; 2000US-0251518P.
 XX PA (CANADA) NAT RES COUNCIL CANADA.
 XX PI Krochko JE, Cutler AJ, Abrams SR;
 XX PI WPI: 2002-519663/55.
 XX DR N-PSDB; ABKS0986.
 XX PT New isolated and purified DNA that encodes protein having abscissic acid
 XX PT (ABA) hydroxylase activity, useful for altering catabolism of abscissic
 XX PT acid in plants.
 XX PS Disclosure: Page 108-111; 117pp; English.
 XX CC The invention relates to an isolated and purified DNA (I) that encodes a
 XX CC protein having abscissic acid (ABA) hydroxylase activity. (I) is useful
 XX CC for producing a transgenic plant which involves introducing (I) into a
 XX CC genome of the plant or its part, and carrying out plant growth and
 XX CC development. (I) is useful for modifying catabolism of ABA or ABA
 XX CC analogues in a plant. ABA hydroxylase clone pB10-30-3 cDNA sequence is
 XX CC useful for identification of related sequences from other plant species.
 XX CC (I) is also useful for altering ABA and ABA analog catabolism in plants.
 XX CC The present sequence represents abscissic acid (ABA) hydroxylase CYP72A14
 XX CC

SO Sequence 512 AA;
 Query Match 39.9%; Score 1089.5; DB 5; Length 512;
 Best Local Similarity 40.7%; Pred. No. 1.9e-97;
 Matches 211; Conservative 114; Mismatches 171; Indels 23; Gaps 6;

QY 12 KYVLVSYSLSYIV-----KQMSLWMPRKIEHESKQIGRGPYHPTGNVKELVG 64
 DB 2 EISVSSVYFSLAVVAVVSWMTLKVWFTPKMLERSLRQGLSTSTPLIGPKKMS 61
 QY 65 MMUKASSHMPSPSHNLPRLVLSFYHMKIKYGAFLVWFGPTFRLLTVADPDLIREISKS 124
 DB 62 MFLKATSKIRKTDITFRVWHP-LQMLKTHGRITLWFGPIPTITMDPOIKKEVFNKV 121

```
QY 125 EPEYKNAHPLVKQJEGDGLSLKGEKNAHHRKIIISPTHEHNIKLLVPVYLKSVTMDV 184
    :::::::::::::::::::::
Db 122 YDFQKHTFFELSKII-GRGLVSYDDKMAQHRRILINPAFLIEKIKMWHVHFHSCSLIV 180
    :::::::::::::::::::::
QY 185 KMSDKLSENG-EVEVDVVEWFPILTEDVIISRTAFGSSYEDGRAVFRLOAQOMLLCAEAFQ 243
    :::::::::::::::::::::
Db 181 EMDKLVSDKSGSEVDVWFGILSMWTADVISRTAFGSSYREGHRIFFELQAEIQAQVMQAFQ 240
    :::::::::::::::::::::
QY 244 KVFIPGIRPPPTKGNLKSRLDKETKRSLLKIEHRRONALDGGEGEKEAQAQDLIGLM 303
    :::::::::::::::::::::
Db 241 KEFIPGYIYLPKGNRMKTAREIODILRGINKRER-----ARSGEAPSDDLIGIL 294
    :::::::::::::::::::::
QY 304 IQA-----KNTVODIVEECKSFFFAKQTTSNLITMTTILISMPEWQAKARDEVLR 356
    :::::::::::::::::::::
Db 295 IBSNLQDTGNGNSTEDWMECKLFFLAGQETTSVLVWTVLISQHDQARAREVYKQ 354
    :::::::::::::::::::::
QY 357 VCGSRDVPTRKDHVVKLKITSLMILNESLRLYPPIVATIRASDVYLGQYKIPCTELLIP 416
    :::::::::::::::::::::
Db 355 VFQDKQ-PDTGGLNQLKWTMTLYEVLRLYPEVQLTRAHKEKLGDLTLPGVQISLP 413
    :::::::::::::::::::::
QY 417 IIVVHDDQIMGNDVNEFPAPFADGVPRRAKHPIGFIPIFGVGTICGMALILQAKLT 476
    :::::::::::::::::::::
Db 414 VLVHRDTLWNGDAGEFKPEERFKDJSKATNNOVSFFPFAMGRICIGQNFILLKXMA 473
    :::::::::::::::::::::
QY 477 LAVMIQRTFHLAPYQHAFTVLMILYPOHGAFTFRRL 515
    :::::::::::::::::::::
Db 474 MSLLQRFSELPSSYVHAHYITITLYPQGAHMLMHL 512
    :::::::::::::::::::::

RESULT 5
AAG31392
ID AAG31392 standard; protein; 518 AA.
XX
XX AAG31392;
AC
XX
XX 17-OCT-2000 (first entry)
Dt
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 37691.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PR
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132866P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136329P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138549P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 15-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
```

PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145293P.
 PR 27-JUL-1999; 99US-01452918P.
 PR 27-JUL-1999; 99US-01452919P.
 PR 28-JUL-1999; 99US-0145291P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148312P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152362P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155132P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158022P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.

PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 34.5%; Score 943; DB 3; Length 518;
 Best Local Similarity 37.5%; Pred. No. 4,4e-83;
 Matches 196; Conservative 107; Mismatches 195; Indels 24; Gaps 8;

QY 10 IPKVLVSLVLSIVT---KGNLSLWREPRKIEHFRSKQIGRPYPHFIGNVKSLVGM 65
 DB 4 ISTNLITVLLEFVSVKIKACMILLRLPMLSKFKQGI6PKYKILYGNLSIKKM 63
 QY 66 MLKSSHPM-PFSNIIIPRVLSFYHMKRYGATFLVWGQFTRLTVADDLRELF-SK 123
 DB 64 KKEADLCVLDPNNSNDIPRVFPQYHQM5QYGTFLFWGTGKPTIYISNHELAKQVLSK 123
 QY 124 SEFEKNEAHPVYKQLEGDGLSLKGEKMAHNRKIISPTFHEMLKLLVAVVLYKSVTDVY 183
 DB 124 FGFITIVKRPVETILGKGLSFTQGDWIRHRILNPAFSMDRLKAMQPMODCLRIF 183
 QY 184 DKMSDKLSENGEV--EVDVYEWFOITEDVYSRTAFSSYEDGRAVFRIOQOMLCAEA 241
 DB 184 EEMR-KORNGEVLIKIEISKRFHKLADIATTAAGSSYAEIEICRSQTELEKYIIS 242
 QY 242 FQKVFIPGVRFPFRGVLKSKRLDKE-RKSLILERRRONADGGECKEPA-AKDLL 300
 DB 243 LTNVFIPTGYLPPLNKLMEHLKKVKSIRKILDSRLKS-----KCKTYGYDDLL 295
 QY 301 GLMIQA-----KNVTVDIVECKSFPPAGKQTSNLLTWTTILSNHEPMQAKARDE 353
 DB 296 GVMULTAKSNVEYKHEMDEIIIECKNFYAGGTSILLTWTTMLSLHQVQKTLREB 355
 QY 354 VLRVGSRDVPTKDHVYKLTLSMLNLSKALPPIVATIRRAKSDVYKGYVPCGTEI 413
 DB 356 VFNECGKDKLPDITFESKLKLMNVLMESIRLGPYKISREKTDQMKVGHLEIPKGTGI 415
 QY 414 LPIIYAVHDOALWGNVNEFPNPARFADGVPRAKHPVGFIPGLGVRTICGNLAILQA 473
 DB 416 IIPILKXHRDKALIGEDAEQFNPLRFENGISQATIHNNLLPSPISGPACIANKFMVBA 475
 QY 474 KLLTAVMIGRFTFLAFTYQHAFTVLMYLPQGAIPTRRL 515
 DB 476 KTVLTMILQOFOLSLSEPKYHTPYDHFDPYQGLPVMHPL 517

RESULT 6
 ADA48312
 ID ADA48312 standard; protein; 454 AA.
 ADA48312;
 AC 20-NOV-2003 (first entry)
 DT XX
 DE Rice protein conferring disease resistance in plants.
 KW disease resistance; pathogen tolerance; plant pathogen; plant; rice.
 OS Oryza sativa.
 OS WO200300906-A2.
 PN 03-JAN-2003.
 PD 21-JUN-2002; 2002W0-IB002453.
 XX
 PF

XX 22-JUN-2001; 2001US-0300112P.
 PR 26-SEP-2001; 2001US-0352277P.
 PR 22-MAR-2002; 2002US-036535P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;
 PI Katagiri F, Kreps J, Provart N, Rieke D, Zhu T;
 XX
 DR WPI; 2003-184052/18.
 DR N-PSDB; ADA48311.
 XX
 PT New polynucleotide comprising a plant nucleotide sequence having an open
 PT reading frame that encodes a polypeptide associated with disease
 PT resistance, useful for conferring resistance or tolerance to a plant
 PT pathogen.
 XX
 PS Claim 10; SEQ ID NO 382; 299pp; English.
 XX
 CC The invention relates to a novel isolated polynucleotide comprising a
 CC plant nucleotide sequence having an open reading frame that encodes a
 CC polypeptide associated with disease resistance or its fragment having
 CC substantially the same activity as the full-length polypeptide. The
 CC polynucleotide of the invention is useful for conferring resistance or
 CC tolerance to a plant pathogen. The present sequence represents a protein
 CC conferring disease resistance used in the invention.
 CC
 SQ Sequence 454 AA;
 XX
 Query Match 33.8%; Score 921.5; DB 6; Length 454;
 Best Local Similarity 41.9%; Pred. No. 4.7e-81;
 Matches 181; Conservative 88; Mismatches 142; Indels 21; Gaps 6;
 XX
 QY 96 GATFLVWFGPTFLVADPDLREIFS-KSEFEKNEAHLVKOLEGGLSLKGEKMAH 154
 DB 30 GRFFLYWFGAQNICLADVSMQVLSDRGTGYPKLTNPHFVRLGSLVLTDDDEK 89
 QY 155 HRRKISPTFEMENLKLVVPLVLSVTDVYKNSDKL-SENGEVVDVYEMFQILTEVVIS 213
 DB 90 HRRVYHFAFMIDLKMTMTMSPDCSRSMSESESLAKGGLVEIELSRREFELTADVIS 149
 QY 214 RTAFGSSYEDGRAVFLQAOQMLLCEAFQKPIFGYRFFPRGNLKSRLDKETRSLL 273
 DB 150 HTAFGSSYKGGKQVFLAQRLEPLAFSTFLVQIPGFSYLPFMKPKFKWSLDKVRGMLM 209
 QY 274 KLIERRRQNA-IDEGECKEPPAKDILGLMTQA-----KNVTVODIVECKSPFF 323
 DB 210 DIKTRANKDVAGYG-----NDLLGLMLEACAPHEHGDSCPLSDDEIIDEKTFEF 261
 QY 324 AGKQTSNLLTWTTILISNHPMQAARDVLRVCGSRDVPYTKHVVYKLTSLINSL 383
 DB 262 AGHDITSHLLTWTFILSTHPDQEKLRBIAMECDK-VPTGDMNLNKLKMNMLLETTL 320
 QY 384 RLVPPIVATIRRAKSDVKGKYIKPGTELLPIIVAHDOAIMGNDVNEFPARFADGV 443
 DB 321 RLISPVSLIRKVDTDIELGIGIMPAGALLTPIATIHDXWGEDADEPDEFKENGCV 380
 QY 444 PRAAKHPVGFIPFGLGVRTICIGNLAILQAKTLTAVMIOFPFPHLAPYQAAVYTMLLY 503
 DB 381 TRAAKHNNALISFSSGPRSCIGQNFAMIEKAVIAMILQRFSTLSPKVNAPDTVITLR 440
 QY 504 PQHGAPITFRRL 515
 DB 441 PKIGLPIVILKSL 452

RESULT 7
 ADA48676
 ID ADA48676 standard; protein; 458 AA.
 AC ADA48676;
 XX

DT 20-NOV-2003 (first entry)
 XX
 DE Rice protein conferring disease resistance in plants.
 XX
 KW disease resistance; pathogen tolerance; plant pathogen; plant; rice.
 XX
 OS Oryza sativa.
 XX
 PN MO2003000906-42.
 XX
 PD 03-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-1B002453.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 PR 26-SEP-2001; 2001US-0352277P.
 PR 22-MAR-2002; 2002US-036535P.
 XX
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;
 PI Katagiri F, Kreps J, Provart N, Rieke D, Zhu T;
 XX
 DR WPI; 2003-184052/18.
 DR N-PSDB; ADA48675.
 XX
 XX
 PT New polynucleotide comprising a plant nucleotide sequence having an open
 PT reading frame that encodes a polypeptide associated with disease
 PT resistance, useful for conferring resistance or tolerance to a plant
 PT pathogen.
 XX
 PS Claim 10; SEQ ID NO 746; 299pp; English.
 XX
 CC The invention relates to a novel isolated polynucleotide comprising a
 CC plant nucleotide sequence having an open reading frame that encodes a
 CC polypeptide associated with disease resistance or its fragment having
 CC substantially the same activity as the full-length polypeptide. The
 CC polynucleotide of the invention is useful for conferring resistance or
 CC tolerance to a plant pathogen. The present sequence represents a protein
 CC conferring disease resistance used in the invention.
 CC
 SQ Sequence 458 AA;
 XX
 Query Match 33.7%; Score 919.5; DB 6; Length 458;
 Best Local Similarity 39.5%; Pred. No. 7.4e-81;
 Matches 188; Conservative 90; Mismatches 155; Indels 43; Gaps 8;
 XX
 QY 50 PYYHFFIGNVKEL-----VGMMLKASHMPFSNHLIPVLSFYHHMKIYATFLVW 102
 DB 14 PYVQFPSSGNLGIKXFRGDAQVNLNV-----SHDPLPIVQPHFRKPIYLRTEFLYW 67
 QY 103 FGFTRFLVADPDLREIFS-KSEFEKNEAHLVKOLEGGLSLKGEKMAHRRKTIISP 161
 DB 68 FGAQPNICLADVSMQVLSDRGTGYPKLTNPHFVRLGSLVLTDDDEKRRKRVYHP 127
 QY 162 TFHEENLKLVVPLVLSVTDVYKNSDKL-SENGEVVDVYEMFQILTEVISTAFRGSS 220
 DB 128 AFNMDLTKMTMTMSPDCSRSMSESESLAKGGLVEIELSRREFELTADVISHTARSS 187
 QY 221 YEDGRAVFLQAOQMLLCEAFQKPIFGYRFFPRGNLKSRLDKETRSLLKLIERR 280
 DB 188 YKGGKQVFLAQRLEPLAFSTFLVQIPGFSYLPFMKPKFKWSLDKVRGMLMILIKRH 247
 QY 281 QNA-IDEGECKEPPAKDILGLMTQAQNVTVODIVECKSPFPAKQTSNLLTWTTIL 339
 DB 248 ANKDVAAGYG-----NDLLGLMLEAMH-----DTTSHLLTWTFIL 281
 QY 340 LSNHPMQAARDVLRVCGSRDVPYTKHVVYKLTSLINSLRLVPPIVATIRRAKSD 399
 DB 282 LSTHPMQEKLRBIAMECDK-VPTGDMNLNKLKMNMLLETTLRVSPVSLIRKVDTD 340
 QY 400 VKLGKYIKPGTELLPIIVAHDOAIMGNDVNEFPARFADGVRAAKHPVGFIPFGLG 459

```
DB      341  IEIGIGIMPEGALLTPIATIRHDKVEWGEDADEFRPERFENGVTAAAGHPNALLSFSAG 400
OY      460  VRTTIGONAILQAKTLTAMVIOQFTPHLAPTQAHAPVTMLTYPOHAPITFRLL 515
DB      401  PRSCIGONFAMIEKAVIAMILORFSPTLSPKYVHAPTDVITLPRKGLPMILKSL 456

RESULT 8
AAG31393
ID      AAG31393 standard; protein: 484 AA.
AC      AAG31393;
XX      17-OCT-2000 (first entry)
DT      17-OCT-2000 (first entry)
XX      Arabidopsis thaliana protein fragment SEQ ID NO: 37692.
DE      Arabidopsis thaliana protein fragment SEQ ID NO: 37692.
XX      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
XX      termination sequence.
OS      Arabidopsis thaliana.
XX      EP1033405-A2.
XX      06-SEP-2000.
XX      25-FEB-2000; 2000EP-00301439.
XX      25-FEB-1999; 99US-011825P.
PR      05-MAR-1999; 99US-0133180P.
PR      09-MAR-1999; 99US-0123548P.
PR      23-MAR-1999; 99US-0125788P.
PR      25-MAR-1999; 99US-0126264P.
PR      29-MAR-1999; 99US-0126785P.
PR      01-APR-1999; 99US-0127462P.
PR      08-APR-1999; 99US-0128232P.
PR      16-APR-1999; 99US-0128714P.
PR      19-APR-1999; 99US-0129845P.
PR      21-APR-1999; 99US-0130077P.
PR      23-APR-1999; 99US-0130449P.
PR      23-APR-1999; 99US-0130510P.
PR      23-APR-1999; 99US-0130891P.
PR      28-APR-1999; 99US-0131448P.
PR      30-APR-1999; 99US-0132048P.
PR      04-MAY-1999; 99US-0132407P.
PR      05-MAY-1999; 99US-0132484P.
PR      06-MAY-1999; 99US-0132466P.
PR      06-MAY-1999; 99US-0132487P.
PR      07-MAY-1999; 99US-0132867P.
PR      11-MAY-1999; 99US-0134256P.
PR      14-MAY-1999; 99US-0134218P.
PR      14-MAY-1999; 99US-0134219P.
PR      14-MAY-1999; 99US-0134221P.
PR      14-MAY-1999; 99US-0134370P.
PR      18-MAY-1999; 99US-0134768P.
PR      19-MAY-1999; 99US-0135941P.
PR      20-MAY-1999; 99US-0135124P.
PR      21-MAY-1999; 99US-0135353P.
PR      24-MAY-1999; 99US-0135629P.
PR      25-MAY-1999; 99US-0136021P.
PR      27-MAY-1999; 99US-0136392P.
PR      28-MAY-1999; 99US-0136782P.
PR      01-JUN-1999; 99US-0137222P.
PR      03-JUN-1999; 99US-0137502P.
PR      04-JUN-1999; 99US-0137528P.
PR      07-JUN-1999; 99US-0137724P.
PR      08-JUN-1999; 99US-0138094P.
PR      10-JUN-1999; 99US-0138540P.
PR      10-JUN-1999; 99US-0138647P.
PR      14-JUN-1999; 99US-0139119P.
PR      16-JUN-1999; 99US-0139452P.
PR      16-JUN-1999; 99US-0139453P.
PR      17-JUN-1999; 99US-0139492P.
PR      18-JUN-1999; 99US-0139454P.
PR      18-JUN-1999; 99US-0139455P.
PR      18-JUN-1999; 99US-0139456P.
PR      18-JUN-1999; 99US-0139457P.
PR      18-JUN-1999; 99US-0139458P.
PR      18-JUN-1999; 99US-0139459P.
PR      18-JUN-1999; 99US-0139460P.
PR      18-JUN-1999; 99US-0139461P.
PR      18-JUN-1999; 99US-0139462P.
PR      18-JUN-1999; 99US-0139463P.
PR      18-JUN-1999; 99US-0139750P.
PR      18-JUN-1999; 99US-0139763P.
PR      21-JUN-1999; 99US-0139817P.
PR      22-JUN-1999; 99US-0139899P.
PR      23-JUN-1999; 99US-0140353P.
PR      23-JUN-1999; 99US-0140354P.
PR      24-JUN-1999; 99US-0140695P.
PR      28-JUN-1999; 99US-0140823P.
PR      29-JUN-1999; 99US-0140991P.
PR      30-JUN-1999; 99US-0141287P.
PR      01-JUL-1999; 99US-0141842P.
PR      02-JUL-1999; 99US-0142055P.
PR      06-JUL-1999; 99US-0142390P.
PR      08-JUL-1999; 99US-0142803P.
PR      09-JUL-1999; 99US-0142920P.
PR      12-JUL-1999; 99US-0142977P.
PR      13-JUL-1999; 99US-0143542P.
PR      14-JUL-1999; 99US-0143624P.
PR      15-JUL-1999; 99US-0144005P.
PR      16-JUL-1999; 99US-0144085P.
PR      16-JUL-1999; 99US-0144086P.
PR      19-JUL-1999; 99US-0144325P.
PR      19-JUL-1999; 99US-0144331P.
PR      19-JUL-1999; 99US-0144332P.
PR      19-JUL-1999; 99US-0144333P.
PR      19-JUL-1999; 99US-0144334P.
PR      19-JUL-1999; 99US-0144335P.
PR      20-JUL-1999; 99US-0144352P.
PR      20-JUL-1999; 99US-0144632P.
PR      21-JUL-1999; 99US-0144814P.
PR      21-JUL-1999; 99US-0145086P.
PR      21-JUL-1999; 99US-0145088P.
PR      22-JUL-1999; 99US-0145089P.
PR      22-JUL-1999; 99US-0145087P.
PR      22-JUL-1999; 99US-0145089P.
PR      22-JUL-1999; 99US-0145192P.
PR      23-JUL-1999; 99US-0145145P.
PR      23-JUL-1999; 99US-0145218P.
PR      23-JUL-1999; 99US-0145224P.
PR      26-JUL-1999; 99US-0145276P.
PR      27-JUL-1999; 99US-0145913P.
PR      27-JUL-1999; 99US-0145918P.
PR      27-JUL-1999; 99US-0145919P.
PR      28-JUL-1999; 99US-0145951P.
PR      02-AUG-1999; 99US-0146386P.
PR      02-AUG-1999; 99US-0146388P.
PR      02-AUG-1999; 99US-0146389P.
PR      03-AUG-1999; 99US-0147038P.
PR      04-AUG-1999; 99US-0147204P.
PR      04-AUG-1999; 99US-0147302P.
PR      05-AUG-1999; 99US-0147260P.
PR      06-AUG-1999; 99US-0147303P.
PR      06-AUG-1999; 99US-0147416P.
PR      09-AUG-1999; 99US-0147453P.
PR      09-AUG-1999; 99US-0147935P.
PR      10-AUG-1999; 99US-0148171P.
PR      11-AUG-1999; 99US-0148319P.
PR      12-AUG-1999; 99US-0148341P.
```

PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149922P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156559P.
PR 28-SEP-1999; 99US-0156588P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.6%; Score 917; DB 3; Length 484;
Best Local Similarity 38.2%; Pred. No. 1,4e-80;
Matches 187; Conservative 100; Mismatches 183; Indels 20; Gaps 7;

QY 38 IEHFSSKQIGRPFYHFFIGNVKELVGMMLKASSHPM-PPSHNLLPRVLGFYHWRKJYG 96
DB 2 LSKPFKQGISGPKYKLLYGNLSFKKKKKKADLCVLDPSNNDIFPRVFPQYHQMWSQYG 61
QY 97 ATFLVWTFPTFRILTAVADPDILIRELF-SKSEFYKRNAAHPLVYQLEDDGLSLKGGKMAHH 155

DB 62 DTFPLWGTQKPTLYISNHELAKOVLSKFGFTIIPVKRDEVEFLIFGKGSLFLOGDMWRH 121
QY 156 RKIISPTFHENLKLIVPVVLKSVTDVNDKWSKLSSENGEV--EVDVVEWFOILFEDVIS 213
DB 122 RRIINPAFSDRLKAMQFPGDCTLRIFEEWR-KQRRNGEVLIKIISKEFKLITDILA 180
QY 214 RTAFSSYEDGRAVFRILQAOOMLLCAEAFQVFIIPGYRFFPTGNKSRKLDKXIRKSL 273
DB 181 TTAFGSSAEBIEELCRQTELEKXYISLNVFIPGTQYLPFTNKLKWLHKKVNGSIRK 240
QY 274 KLIERRONAIDGGECEKERA-AKDLIGLMIA-----KNVTQDIYECKSPFFAG 325
DB 241 RIIDSRKLS-----KCKTYGYGDDLLGYMLTAASNEYERKMRDEILIECKNFTYAG 293
QY 326 KQTSNLLTWTTLISNHPWQAKARDEVLRVCGSRDVPYKDHVYKLTLSMLTNSLR 385
DB 294 QGTISILLTWTMLSLHQGQEKLRREVNECKKXKIPDTDFSKLKLANVLMESLRL 353
QY 386 YPPIVATIRRAKSDVKLGKIKIPCGTELLIPITAVHDDQAINGVNDVNEFRAPFAGV 445
DB 354 YGPVAKISRERATQDMKYGHLEIPKGTIILPLMKHDKAIWGEDAEQFPLFEENGISO 413
QY 446 AAKHVGFIPRGVATCGIGONLAILOAKLTLAWIQRFPHLAPTYOHAPTYMLMLYPQ 505
DB 414 ATIHFNALLPFSIGPRACIAKNFMWAKVLMILQFQSLSPFEKHPVDFHFLDFPQ 473
QY 506 HGAPITFRRL 515
DB 474 YGLPVMLEPL 483

RESULT 9
ID AAG31394
AAG31394 standard; protein; 456 AA.
XX
AC AAG31394;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37693.
XX
DE Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 25-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134421P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138034P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139112P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139482P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139859P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140659P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144844P.
PR 20-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 04-AUG-1999; 99US-0147322P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147939P.
PR 09-AUG-1999; 99US-0148171P.
PR 10-AUG-1999; 99US-0148319P.
PR 11-AUG-1999; 99US-0148341P.
PR 12-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150066P.
PR 26-AUG-1999; 99US-0150084P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157665P.
PR 07-OCT-1999; 99US-0158093P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.

```

PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160982P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

```

Query Match 31.7%; Score 864.5; DB 3; Length 456;
 Best Local Similarity 38.5%; Pred. No. 1,8e-75;
 Matches 174; Conservative 94; Mismatches 165; Indels 19; Gaps 6;

```

QY 75 PESHNLPRLVSLFYHWRKIYGATFLVFGFTFLTVADPDLIREIF-SKSEFYKNEAH 133
DB 12 PMSNDIFRPFVPOYHOMNSQYDFTLFWTGKPTIYISNHELAKQVLSKKGFTIIPVKR 71
QY 134 FLVKQLESGDGLSLKGEKAAHHRKTIISFTFMENIKLVVYVLSVTDMDKMSDKSEN 193
DB 72 PEVFILFGKLSFIOGDDIRRRRLINPAFSMDRLKAMTOPMGDCTLRIFEEMR-KORRN 130
QY 194 GEV--EVDVYEWFOILTEDEVISRTAFSGSYEDGSAVFRLQAOQMLCAEAFQKVFIPGYR 251
DB 131 GEVLKIRSKFHLTDITFTAFSGSYAEGIELCSQLEKVIYISLTVNFIPTQ 130
QY 252 PPTPGNLSKRLDKIRKSLIKLIERRONALIDGEGCEKPA-ADULGLMIAQ--- 306
DB 191 YLPTPTNLKWLHKKVNSIRIIDSRLKS-----KCTIYGYGDLGLVMTAKSN 243
QY 307 ---KNTVGDIVIECKSPFAGKOTSNLTTTILSMHEMOAKADEVLKVGSGRDV 363
DB 244 EYERKRRMDIIECKNFYFAGGTSILTTTILSLHGMQEKREBFVFNCGDKI 303
QY 364 PTKDHVYKLTSLMILNESLRYPPIVATIRAKSDVKLGKIKIPGTELLPIIAVHD 423
DB 304 PPDTRFSKLLMMVMYMSLRLYGPVAKISREATQDMKVLHBPKGTSIIIPILKQHRD 363
QY 424 QAIWMDVNEFNPAPRADGVPRAAHPVCFIPFGVGTCTIGOMLALIAQAKLTAVNIQR 483
DB 364 KAIWGEDAQFNPLRENGISQATHPNALPFSISGPRACIAKQFAMVEAKTVLMTILQQ 423
QY 484 FTFHLAPTYQHAFTVLMLLYPOHGAFTTFRRL 515
DB 424 FQLSLSPRYKHTPVDFDLFPQYGLFVMLHPL 455

```

RESULT 10
 ABJ05536
 ID ABJ05536 standard; protein; 505 AA.

XX ABJ05536;
 DT 14-NOV-2002 (first entry)

XX Breast cancer-associated protein 1.
 XX Breast cancer; breast cancer-associated gene sequence; drug development;
 KM pharmacogenetics; biosensor development.
 XX Unidentified.

XX OS
 XX NC0200259377-A2.
 XX PN
 XX 01-AUG-2002.
 XX PD
 XX 24-JAN-2002; 2002MCO-US002242.

```

XX 24-JAN-2001; 2001US-0263965P.
XX 02-FEB-2001; 2001US-0265928P.
XX 09-APR-2001; 2001US-00829472.
XX 09-APR-2001; 2001US-0282658P.
XX 04-MAY-2001; 2001US-0288590P.
XX 29-MAY-2001; 2001US-0294443P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Mack DH, Gish KC, Afar D;
XX WPI; 2002-583738/62.
XX N-PSDB; ABT07693.
XX
XX Detecting a breast cancer-associated transcript in a patient's cell,
XX useful for diagnosing breast cancer, comprises contacting a biological
XX sample with a polynucleotide that selectively hybridizes with breast
XX cancer nucleic acids.
XX
XX Disclosure; Page 347; 414pp; English.
XX
XX The invention comprises a method of detecting a breast cancer-associated
XX transcript in a cell from a patient. The method of the invention involves
XX contacting a biological sample from the patient with a nucleotide that
XX hybridizes to one of the 69 breast cancer-associated gene sequences shown
XX in the specification. The method of the invention is useful in the
XX diagnosis or prognosis of breast cancer, and for detecting genes that are
XX up or down-regulated in breast cancer cells. Genes identified by the
XX method of the invention can be used in diagnostic purposes and also as
XX targets for screening for therapeutic compounds that modulate breast
XX cancer (e.g. hormones or antibodies). Identification of genes that are
XX over or under expressed in breast cancer can additionally provide high-
XX resolution, high-sensitivity datasets which can be used in the areas of
XX diagnostics, therapeutics, drug development, pharmacogenetics, protein
XX structure and biosensor development. Amino acid sequences ABJ05536 -
XX ABJ05604 represent the proteins encoded by the 69 breast cancer-
XX associated genes of the invention
XX
XX Sequence 505 AA;

```

Query Match 19.9%; Score 543.5; DB 5; Length 505;
 Best Local Similarity 29.3%; Pred. No. 6.1e-44;
 Matches 149; Conservative 104; Mismatches 155; Indels 61; Gaps 20;

```

QY 11 PAVLVSYLSLVYKGSMLIM--WRPKIEHFSKQGRGPYHFFIGNVKELVGMML 67
DB 13 PVLILILCMISLLLEQVIRLYQRRRMIRAL--HL-----PPAPAHVFYGH-KEF----- 60
QY 68 KASSHPMPSPSHNILERVLSFYHWRKIYGATFLVMBP--TPRLTVADPDLIREIFSKSEF 126
DB 61 ----YPV-----KEPEVYHKLMERYPCAVPLMVGPFTMFPSVHPDPAK--ILLKQD 107
QY 127 YEKNAHPVPLVQLEGDGLSLKGEKAAHHRKTIISFTFMENIKLVVYVLSVTDMVQX 186
DB 108 PKSAVSHKILSWRGVLTLDGSKMKKHQIVKPGFNISILKFTIMGESVYMLNKKV 167
QY 187 SDKLSNGEVEVDVYEMFQILTEDEVISRTAF---GSSYEDG-----RAYFRQA---QQ 234
DB 168 EERIAQNSRLK--LFQVSLMTLDSIMKCAFSHQSGSIOBSTDLSYKAAVFNLSKISNQR 225
QY 235 MLCALAFQKVFIPGYRFPFRGNLSKRLDKERKSLKLIERRRONAVIDGEECKEP 294
DB 226 MNMPLHNDLVF-----KFSOGQILFS-KENQELHQTEKVIQDRKESLKDIXQDPTQK 279
QY 295 AAKDILGLMIQAKNTVDIVF-----ECSFFPAGKOTTSNLTWTITLLSMPEMOK 349
DB 280 RRMDFDLILLSAKSENTDFSEADLQAEVKTFMFGHDTTSSAISWLYCLAKYBEHQOR 339
QY 350 ARDEVLYKVGSGRDVPTKHVVKLTSLMILNESRLVPIVATIRAKSDVKL-GGYXIP 408
DB 340 CRDEIRELLDGSSITWHLISOMPTTMCICELRIYAFVNISRLLDKDPIITPFGSGSLP 399

```

QY 409 CGTELLIFIIAVHDDQAIWGVNPFARFADGVPPAAK-HVGVIFPGLGYRTICGN 467
DB 400 AGITVFIMIALHNHPFW-EDPOVFNPLFRSR--ENSEKIHPPAFIPFAGJRNICIGH 456
QY 468 LALIQAUKTLAVMIQRTFPHLAPYQAP 496
DB 457 FALTECKVAVALTLIR--FKLAPDHSRPP 483

RESULT 11
AAE14447
ID AAE14447 standard; protein; 505 AA.
AC AAE14447;
XX
XX
XX 26-MAR-2002 (first entry)
DE Human drug metabolizing enzyme (DME)-10.
XX
XX Human; drug metabolizing enzyme; DME-10; autoimmune; inflammatory;
XX cell proliferative; developmental; endocrine; eye; metabolic; AIDS;
XX gastrointestinal disorder; liver disorder; cancer; arteriosclerosis;
XX adult respiratory distress syndrome; anaemia; epilepsy; hypothyroidism;
XX hypochlasmia; pituitary; diabetes; hypogonadism; conjunctivitis;
XX glaucoma; cystic fibrosis; hypercholesterolaemia; gastritis;
XX peptic ulcer; hepatitis; gene therapy.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX Peptide 1..31
XX /label= Signal_peptide
XX /label= 33
XX Domain /label= Transmembrane_domain
XX /label= 32..505
XX Protein /label= Mature_DME-10
XX
XX WO200190334-A2.
XX
XX
XX 29-NOV-2001.
XX
XX
XX 25-MAY-2001; 2001WO-US017150.
XX
XX
XX 25-MAY-2000; 2000US-0207901P.
XX 01-JUN-2000; 2000US-0208983P.
XX 07-JUN-2000; 2000US-0209861P.
XX 15-JUN-2000; 2000US-0211825P.
XX 22-JUN-2000; 2000US-0213744P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Sanjanwala VS, Baughn MR, Gandhi AR, Ring HZ, Elliott V;
XX Maia NK, Yang J, Khan PA, Ramkumar J, Tang YT, Hafeela A, Lal P;
XX Nguyen DB, Yeo MG, Lee EA, Tribouley CM, Patterson C, Lu Y,
XX Burford N, Ding L, Bruns CM, Kearney L, Reddy R;
XX
XX MPI; 2002-097650/13.
XX N-PSDB; AAD24015.
XX
XX New human drug metabolizing enzymes and polynucleotides encoding the
XX enzyme for diagnosing, preventing or treating cell proliferative,
XX autoimmune/inflammatory, endocrine, eye, metabolic and gastrointestinal
XX disorders.
XX
XX Claim 1; Page 145-146; 158pp; English.
XX
XX The present sequence is human drug metabolizing enzyme (DME)-10. DME
XX polypeptide, polynucleotide and modulators are useful for diagnosis,
XX treatment and prevention of autoimmune/inflammatory, cell proliferative,
XX developmental, endocrine, eye, metabolic, and gastrointestinal disorders,
XX including liver disorders. The autoimmune/inflammatory disorders
XX treatable include AIDS, adult respiratory distress syndrome, Addison's
XX disease, allergies, anaemia, asthma, arteriosclerosis, osteoporosis,

CC autoimmune hemolytic anaemia, autoimmune thyroiditis, Crohn's disease,
CC atopic dermatitis, diabetic mellitus, Graves' disease, systemic lupus
CC glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus
CC erythematosis, systemic sclerosis, ulcerative colitis, haemodialysis and
CC urethritis, viral, bacterial, fungal, parasitic, protozoal, helminthic
CC infections and trauma, and cell proliferative disorders such as cancer,
CC actinic keratosis, arteriosclerosis, atherosclerosis, bursitis,
CC cirrhosis, hepatitis and psoriasis. Developmental disorders include
CC anaemia, renal tubular acidosis, epilepsy, hypothyroidism and cataract,
CC and endocrine disorders include disorders of hypothalamus and pituitary,
CC disorders associated with hypopituitarism, including sarcoidosis,
CC diabetes insipidus, hypogonadism, disorders associated with
CC hypothyroidism including goitre, acute thyroiditis, Graves' disease,
CC disorders associated with hyperparathyroidism, pancreatic disorders such
CC as type I or type II diabetes mellitus, disorders associated with
CC adrenals such as hyperplasia, Cushing's disease, endometriosis,
CC infertility, hypergonadal disorders, and gynaecomastia. Eye disorders
CC include conjunctivitis, keratitis, glaucoma and macular degeneration, and
CC metabolic disorders include diabetes, cystic fibrosis, goitre,
CC hypercholesterolaemia, hypoglycaemia, hyperhidrosis, lysosomal storage
CC diseases, obesity, phenylketonuria and hypocalcaemia. Also the molecules
CC are useful for treating gastrointestinal disorders such as dysphagia,
CC gastritis, peptic ulcer, cholelithiasis, cirrhosis, hepatitis,
CC hyperbilirubinaemia, constipation, diarrhoea, jaundice, Wilson's disease,
CC thrombosis and hepatic tumours. The DME polypeptide is also useful for
CC screening its agonist or antagonist
XX
XX Sequence 505 AA;
SQ
Query Match 19.9%; Score 543.5; DB 5; Length 505;
Best Local Similarity 29.3%; Pred. No. 6.1e-44;
Matches 149; Conservative 104; Mismatches 195; Indels 61; Gaps 20;
QY 11 PKVLVSLVSLVVKMSLLM--WRPRIEBHSQGRGPPYFFIGNVELVGMML 67
DB 13 PRLTLICWSLPLPVIRLYKORRRWIRL--HL-----FPAPPAWYFVH-XEF----- 60
QY 68 KASHMPMPFSNHLPRVLSFYHNRKLYGATFLWFGP-TFLTVADPDILREIFSKSE 126
DB 61 ----YPV-----KEFEVYHKMEKPCPAVPLWGPFTMFSVHDPDAK-ILTKRQD 107
QY 127 YEKNEAHPVYKQLEGDGLSLKGEKMAHKKIISPTHEMNLKLVVLYKSTYDWDKX 186
DB 108 PKSAVSHKTESWVGRLVTLDSGKWKKHQIVKPGNISILKIFITMSESRYMNLKX 167
QY 187 SDKLSNGEVEVDYEWFOIITEDVLSRTAF---GSSYEDG-----RAVFLQA---QQ 234
DB 168 EERIAQNSRLA--LFGVSLMTLDSIMKCAFSGHQSIOIDSTLDSYKAVFNLSKISNR 225
QY 235 MLCAPAFQKFTIGYAFPPFRGTLKSKKIDKIRKSLTLIRRRQNALDGESECKE 294
DB 226 MNFLHNDLVF---KFSQGGIFS-KNOEIHQTEVYIDRKESLKKDKKQDOTTQK 279
QY 295 AAKDLIGMIQAKVTVODIVE-----ECKSFPFAGKOTTSNLTITLILSNPSPWQAK 349
DB 280 RRMPLDILLAKSENKQDSEADLQAEVTFMFGHDTTSSAISWILYCLAKYPHQQR 339
QY 350 ARDEVLRVCGSRDVPFTDHYVYKLTLSMILNESRLYPPVATIRRAKSVYK-GGYKIP 408
DB 340 CRDEIRELIGGSSITWEHLSQMYTTMCKECLRLAPAVNISRLDKRITPPDGRSLP 399
QY 409 CGTELLIFIIAVHDDQAIWGVNPFARFADGVPPAAK-HVGVIFPGLGYRTICGN 467
DB 400 AGITVFIMIALHNHPFW-EDPOVFNPLFRSR--ENSEKIHPPAFIPFAGJRNICIGH 456
QY 468 LALIQAUKTLAVMIQRTFPHLAPYQAP 496
DB 457 FALTECKVAVALTLIR--FKLAPDHSRPP 483

RESULT 12
AAU83606
ID AAU83606 standard; protein; 505 AA.

XX AAU83606;
AC
XX
DT 08-MAY-2002 (first entry)
XX
DE Human PRO protein, Seq ID No 30.
XX
XX
XX Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
XX breast cancer; prostate tumour; rectal tumour; liver tumour;
KW pericyte cell proliferation; chondrocyte cell proliferation;
KW tumour necrosis factor-alpha.
XX
XX Homo sapiens.
OS
XX
PN WO200208288-A2.
XX
PD 31-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-US021066.
XX
PR 20-JUL-2000; 2000US-0219556P.
PR 25-JUL-2000; 2000US-0220585P.
PR 25-JUL-2000; 2000US-0220605P.
PR 25-JUL-2000; 2000US-0220607P.
PR 25-JUL-2000; 2000US-0220624P.
PR 25-JUL-2000; 2000US-0220638P.
PR 25-JUL-2000; 2000US-0220664P.
PR 25-JUL-2000; 2000US-0220665P.
PR 26-JUL-2000; 2000US-0220893P.
PR 28-JUL-2000; 2000WO-US020710.
PR 01-AUG-2000; 2000US-0222425P.
PR 22-AUG-2000; 2000US-0227133P.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 10-NOV-2000; 2000WO-US030873.
PR 28-NOV-2000; 2000US-0253646P.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-0074725P.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US00666P.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001WO-US017092.
XX
XX (GETH) GENENTECH INC.
PA
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ,
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
PI WPI; 2002-172001/22.
XX
DR N-PSDB; ABK35550.
XX
XX
XX One hundred and twenty two nucleic acids encoding PRO polypeptides,
XX useful for treating a PRO related disorder and for diagnosing tumors such
XX as lung cancer, colon cancer, breast tumor, prostate tumor, rectal tumor
XX or liver tumor.
XX
XX
XX Claim 11, Fig 30, 359pp; English.
XX
XX
XX The invention relates to one hundred and twenty two nucleic acids
XX encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
XX encode human secreted proteins. The PRO nucleic acids, polypeptides,
XX agonists and antagonists are useful for treating a PRO related disorder.
XX The PRO polypeptides are useful for diagnosing tumors, especially lung
XX cancer, colon cancer, breast tumor, prostate tumor, rectal tumor or
XX liver tumor. The PRO polypeptides are useful for stimulating the
XX proliferation of, or gene expression, in pericyte cells, for stimulating
XX the proliferation or differentiation of chondrocyte cells, for
XX stimulating the release of tumour necrosis factor-alpha from human blood,
XX for stimulating or inhibiting the proliferation of normal human dermal
XX fibroblast cells. The PRO polypeptide may also be used as molecular

CC weight markers and for tissue typing. The PRO nucleic acids have
CC applications in molecular biology, including use as hybridisation probes,
CC and in chromosome and gene mapping. AA083592-AA083713 represent human PRO
CC protein sequences of the invention
CX
CX
CX Sequence 505 AA;
CX

Query Match	19.8%;	Score 540.5;	DB 5;	Length 505;
Best Local Similarity	29.3%;	Pred. No. 1.2e-43;		
Matches 149;	Conservative 103;	Mismatches 196;	Indels 61;	Gaps 20;

11 PKVLVSVILSVIKGMSLLW--WRPKIEEHFSKQIGRPYHFFIGNVKELVGMM 67

13 PETLLTLCMSLLFQVIRLYORRRWMIRAL--HL-----FPAPPAHWFYGH-KEF----- 60

12 EKSEFETRTDPIADTVET-EGVYVETATYKBYHSEV DIT TITNINCECIVV 12

[illegible]

D5 61 ----YPV-----KEFEVYHKLMETKPCAVLWGGFIMFSEVDFDIAN-LEDAVQ

127 YEKNEAHPVKQLEGDLLSLKGEKWAHHRKISPTFHMENTKLVPVLLKSVTDMVDKW 128

Db 108 PKSAVSHKILES WVGRLVTL DGS KWK KHRQIVKPGFNISILKIFITMSESVFM LNKW 16

187 SDKLSNGEVEVDVYEMFOILTEDVISRTAF---GSSYEDG---RAVFRQA---QQ 23

[illegible][illegible]

235 MLLCAEAFQKVEIPGIRFFPKGNLNSKNLDNEIANGSLNLAIAVNNNG
:::|:::|:::|:::

Db 226 MNNFLHNDLVF-----KFSOGQIFS-KENQELHQFIEKVIQDRKESLWNNQD : 128

295 AAKDLGLMIQAKNTVQDIVE-----ECKSFFFAKQOTSNLLTWTIILSMHPWQAK 34

Db 280 RRWDFLDILLSAKSENTKDFSEADLQAEVKTFMFAGHDTSSAISWILYCLAKYPEHQQR 33

350 ARDEVLRVCGSRDVP TKDHVVKLKTLSMILNESLR LYPPIVATIRRAKSDVKL-GGYKIP 400

[illegible][illegible]

QY 409 CGIBLLIF1 IAVHHDDAI RONDVINEFNEFNKRDOSV LILBUN III VOIITLII
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Db 400 AGITVFINIWALHNHPYFW-EDPQVENPLRFSK--ENSEKIHPIAFIPISAGLNUVCIGU 1

QY 468 LAILOAKLTIAVMIQRTFHLAPTYQHAP 496

Db 457 FAIECKVAVALTLR--FKLAPDHSRPP 483

RESULT 13

ABU10221
ID ABU10221 standard; protein; 505 AA.

XX
AC
ABU10221:

XX
E
3
F
3
3
3
(
F
3
3
3
3
3
3
3
3
3
)

[illegible]

DE Human cytochrome P450, 3A5.

KW Human; cytochrome; p450; cellular proliferation; cancer; migraine;
KW cellular differentiation; cell adhesion; cell motility; cell migration;

inflammation; autoimmune disorder; rheumatoid arthritis; diabetes; inflammatory bowel disease; conjunctivitis; asthma

KW multiple sclerosis; systemic lupus erythematosus; allergic disease; graft-versus-host disease; vascular disorder;
KW Grave's disease; graft-versus-host disease; vascular disorder;

KW central nervous system disorder; Alzheimer's disease; amnesia; depression; Parkinson's disease; obesity; cachexia; anorexia; KW

cell-signaling-associated disorder; metabolism-associated disorder; cell-signaling-associated disorder; fatty acid-associated disorder.

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	

os
XX
XX
Homo sapiens.

FH	Key	Location/Qualifiers
ET	Periode	1-33

```
FT /label= signal_sequence
```


PT	Protein	34.. .505
FT	/label= Mature_33303	
FT	Domain	46.. .501
FT	/label= Cytochrome_P450_domain	
FT	Modified-site	92.. .95
FT	/note= "Casein kinase II phosphorylation site"	
FT	Modified-site	113.. .115
FT	/note= "Protein kinase C phosphorylation site"	
FT	Modified-site	145.. .148
FT	/note= "N-glycosylation site"	
FT	Modified-site	159.. .161
FT	/note= "Protein kinase C phosphorylation site"	
FT	Modified-site	175.. .178
FT	/note= "Casein kinase II phosphorylation site"	
FT	Modified-site	206.. .209
FT	/note= "Casein kinase II phosphorylation site"	
FT	Modified-site	217.. .220
FT	/note= "N-glycosylation site"	
FT	Modified-site	243.. .248
FT	/note= "N-myristoylation site"	
FT	Modified-site	257.. .259
FT	/note= "Protein kinase C phosphorylation site"	
FT	Modified-site	264.. .267
FT	/note= "cAMP and cGMP-dependent protein kinase phosphorylation site"	
FT	Modified-site	267.. .270
FT	/note= "Casein kinase II phosphorylation site"	
FT	Modified-site	267.. .269
FT	/note= "Protein kinase C phosphorylation site"	
FT	Modified-site	277.. .279
FT	/note= "Protein kinase C phosphorylation site"	
FT	Modified-site	290.. .292
FT	/note= "Protein kinase C phosphorylation site"	
FT	Modified-site	300.. .303
FT	/note= "Casein kinase II phosphorylation site"	
FT	Modified-site	351.. .356
FT	/note= "N-myristoylation site"	
FT	Modified-site	381.. .384
FT	/note= "N-glycosylation site"	
FT	Modified-site	391.. .394
FT	/note= "Casein kinase II phosphorylation site"	
FT	Modified-site	434.. .436
FT	/note= "Protein kinase C phosphorylation site"	
FT	Region	445.. .454
FT	/note= "Cytochrome P450 cysteine haem-iron ligand signature"	
FT	Modified-site	448.. .453
FT	/note= "N-myristoylation site"	
FT	Modified-site	454.. .459
FT	/note= "N-myristoylation site"	
XX		
XX	US2003022334-A1.	
XX	PN	
XX	PD	30-JAN-2003.
XX		
XX	04-FEB-2002; 2002US-00067668.	
XX	PF	
XX	PR	02-FEB-2001; 2001US-0266140P.
XX	PA	(GLUC/) GLUCKSMANN M A.
XX		
XX	Glucksmann MA;	
XX	WI; 2003-447427/42.	
XX	DR	N-PSDB; AChA61906.
XX		
XX	New 33312, 33303 or 32579 nucleic acid molecule encoding cytochrome P450C	
XX	polypeptides, useful for diagnosing and treating cancer, inflammation,	
XX	vacular disorders, CNS disorders or metabolic disorders, and in	
XX	pharmacogenomics.	
XX		
XX	Claim 5, Page 46, 76pp; English.	
XX		

Query	Match	19.8%	Score 540.5	DB 6	Length 505
Best Local Similarity	29.3%	Pred. No. 1.2e-43			
Matches 149	Conservative 103	Mismatches 196	Indels 61	Gaps 20	
QY	11 PKYLVSVILSVIKGMSLIM--WRPKIEEHFSKQIGRPYPHFGVKEVLVGMKL 67				
DB	13 PFLILLILCMGSLLEPOVIRLYQRRRMIRAL--HL-----FPAPPAHVFYGH--KEF----- 60				
QY	68 KASHPMPSPSHILPRVLSFYHMRKLYGATFLWPGP--PRLTVADPDLIEIFSKSEF 126				
DB	61 ----YEV-----KEPEVYHKLMEXKPCAIVLWGPFTMFSSVHPDYAK--ILLRKD 107				
QY	127 YENKKAHPYKQLGEGDGLSLKGEKWAHRRKILSPFHEMNLKLLVYVLLKSVTDVNDXW 186				
DB	108 PKSAVSHKILIESWVGKLVTLDDGSKWKKGRIYVKGNISILKFLTFMSESVRYMLNKM 167				
QY	187 SDKLSNGEVEVDVYEMFOILTEDIYSRTAR--GSSYEDG-----RAVFRQA---QQ 234				
DB	168 EEHIAQNSRL--LPOHYSIKMLTDSIMWCARSHQSGIQDOSTLSDYKAVFNLSKISNR 225				
QY	235 MLCIAEAPQGVFIPGRFPPTGINKSKRKLDEKIRKSLILKILERRONALDGECECKP 294				
DB	226 MANNFLHNDLVF-----KFSGGQGIIFS--KFNQELHQTEKVIQDRKESLKDQKLDQTTOK 279				
QY	295 AAKDLIGLMIOAKNTVDIVF-----ECSFPFAGKQTSNLLTWITLLSHMPRMOK 349				
DB	280 RRMFIDILLAKSEKTKDFSEADLQAEVKTFMGAGDITTSASLWLYCLAKIPRHQR 339				
QY	350 ARDEVLVCGSRVPRKDVVLTXTLSITNESRLVPIVATIRKAKSVKL--GSKYLP 408				
DB	340 CRDRIRELDDGSGITWEHLSQMPPTTCICECRLIAPVYNISRLDKDITPDGRSLP 399				
QY	409 CGTELLPIITVAHDOAIGNDVNEFNPNFADGVPPRAK--HPYGFIPFGLSGRTICGN 467				
DB	400 AGIVTFINIALHNNYFW--EDPOVFNNPLRFSR--ENSEKHIPAYALFISAGLRNCTIGH 456				
QY	468 LAIIQALTLAVVICQRTFHLAPTQHP 496				
DB	457 FALTECVAAVALTLNR--FKLAPDSRRP 483				
RESULT 14					
ID	ABU80753	standard; protein; 505 AA.			
XX	ABU80753;				
XX	23-JUN-2003	(first entry)			
XX	Human PRO	polypeptide #15.			

XX Human; PRO polypeptide; secreted and transmembrane protein;
 KW anti-PRO antibody; diagnostic assay; gene expression; tumour; cytostatic.
 XX
 OS Homo sapiens.
 XX US2003036635-A1.
 XX
 PD 20-FEB-2003.
 XX
 PF 28-AUG-2002; 2002US-00230163.
 XX
 PR 25-JUL-2000; 2000US-0220638P.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-APR-2002; 2002US-00119480.
 XX
 PA (GETH) GENENTECH INC.
 XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
 XX WPI; 2003-342045/32.
 DR N-PSDB; ACA6855.
 XX
 PT One hundred and twenty two nucleic acids encoding PRO polypeptides,
 PT useful for the manufacture of a medicament for diagnosing or treating
 PT tumor.
 PS
 PS Claim 11; Fig 30; 314pp; English.
 XX
 CC The present invention relates to the isolation of novel human PRO
 CC polypeptides, and the polynucleotide sequences encoding them. The PRO
 CC polypeptides are secreted and transmembrane proteins. The PRO
 CC polypeptides and polynucleotides are useful for preparing a medicament
 CC useful in the diagnosis and treatment of tumours. Anti-PRO antibodies are
 CC useful in diagnostic assays for PRO, by detecting its expression in
 CC specific cells, tissues or serum, and for affinity purification of PRO
 CC from recombinant cell culture or natural sources. AB080739-AB080860
 CC represent the human PRO polypeptides of the invention. Note: The sequence
 CC data for this patent was obtained in electronic format directly from the
 CC USPTO web site at seqdata.uspto.gov/patidentry.html
 CC
 CC Sequence 505 AA;
 SQ
 Query Match 19.8%; Score 540.5; DB 6; Length 505;
 Best Local Similarity 29.3%; Pred. No. 1.2e-43;
 Matches 149; Conservative 103; Mismatches 196; Indels 61; Gaps 20;

QY 350 ARDEVLRVCGSRDVFPTDQHVYKLTSLMINSRLRYPPVATIRPAKSDVYL-GGYKIP 408
 Db CRDEIRBELDGGSSITWEHLSQMPYTTMCICEKRLYAPVNVNLSRLDKRITPPDGRSLP 399
 QY 409 CGTELLIFIVVHNDQALMGNVNEFNPAPFADGVPPAAK-HPVGFIPFGAGVTCIGON 467
 Db AGITVFIVTWLHNHPYFW-EDPOVENPLAFSR--ENSEKIHVYALTFPSAGLRNIGQH 456
 QY 468 LAILQAKLTIAWIORPFTFLAPTYOHAP 496
 Db 457 FALIECKVAVALTLIR--FKLAPHSRPP 483
 RESULT 15
 ABO33719
 ID ABO33719 standard; protein; 505 AA.
 XX
 AC ABO33719;
 XX
 DT 17-SEP-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO9862.
 XX
 KW Human; secreted and transmembrane protein; PRO; cytostatic;
 KW antiarthritic; osteopathic; gene therapy; TNF-Agonist-Alpha;
 KW chondrocyte stimulator; pericyte stimulator; fibroblast modulator;
 KW pharmaceutical; diagnostic; biosensor; bioreactor; tumour; lung tumour;
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW liver tumour; bone disorder; cartilage disorder; sports injury;
 KW arthritis; wound.
 XX
 OS Homo sapiens.
 OS
 PN US2003045667-A1.
 XX
 PD 06-MAR-2003.
 XX
 PF 12-AUG-2002; 2002US-00218631.
 XX
 PR 01-JUN-2001; 2001WO-US017800.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-APR-2002; 2002US-00119480.
 XX
 PA (GETH) GENENTECH INC.
 XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
 XX WPI; 2003-512315/48.
 DR N-PSDB; ACDE8607.
 XX
 PT New genes, and its encoded secreted and transmembrane polypeptides,
 PT useful for stimulating Tumor Necrosis Factor alpha, or chondrocyte or
 PT pericyte proliferation, especially for treating lung tumors, arthritis or
 PT wounds in a mammal.
 PS
 PS Claim 11; Fig 30; 314pp; English.
 XX
 CC The invention describes an isolated nucleic acid molecule comprising a
 CC sequence with at least 80% identity to: (a) a nucleotide encoding any of
 CC 122 PRO (secreted and transmembrane) polypeptides whose sequences are
 CC fully defined in the specification; or (b) any of 122 nucleotide
 CC sequences having e.g. 4834, 2504 or 1759 bp fully defined in the
 CC specification; or the full length coding sequence of any these 122
 CC nucleotide sequences. The PRO polypeptides or polynucleotides are useful
 CC as pharmaceuticals, diagnostics, biosensors or bioreactors. These are
 CC particularly useful for detecting tumours (e.g. lung tumour, colon
 CC tumour, breast tumour, prostate tumour, rectal tumour, or liver tumour)
 CC in a mammal, for stimulating the release of TNF-alpha from human blood,
 CC for stimulating proliferation or differentiation of chondrocyte
 CC cells, for stimulating proliferation of pericyte cells, or for modulating
 CC normal human dermal fibroblast proliferation. The PRO nucleic acid or
 CC polypeptide is also useful for treating tumours or various bone and/or

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 10:22:05 ; Search time 21 Seconds
(without alignments)
2381.885 Million cell updates/sec

Title: US-09-992-901-2

Perfect score: 2730
Sequence: 1 MEESSSMFIPKXVLVLSVIL.....LTPQHGAPITRRLLNHEED 520

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: Df11.*
2: Df12.*
3: Df13.*
4: Df14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	2722	99.7	520 2 H84663	probable cytochrome
2	1113	40.8	524 2 T09989	cytochrome P450 -
3	1110	40.7	516 2 T10000	cytochrome P450 (C
4	1101	40.3	524 2 T09944	probable cytochrom
5	1028.5	37.7	505 2 D96781	cytochrome P450, p
6	1023	37.5	508 2 T16980	probable cytochrom
7	1019	37.3	517 2 T02192	cytochrome P450 ho
8	985.5	36.1	526 2 T02191	cytochrome P450 ho
9	943	34.5	518 2 T05876	similar to Cytochr
10	900	33.0	476 2 D86306	hypothetical prote
11	887	32.5	512 2 A96695	laurate omega-hydr
12	523	19.2	509 1 O4RTLO	cytochrome P450 Cy
13	507	18.6	509 2 S47553	hypothetical prote
14	503	18.4	518 2 T24778	hypothetical prote
15	495	18.1	518 2 T24779	hypothetical prote
16	493	18.1	511 1 A34260	laurate omega-hydr
17	488.5	17.9	510 1 A29368	prostaglandin omeg
18	488.5	17.9	518 2 T20908	hypothetical prote
19	487.5	17.9	504 2 A60564	cytochrome P450 3A
20	486.5	17.8	517 2 T09097	hypothetical prote
21	484.5	17.7	519 2 I53015	fatty acid omega-h
22	482.5	17.7	507 2 JCT883	cytochrome P450 4X
23	482	17.7	537 2 JG4534	cytochrome P450 4F
24	481.5	17.6	519 2 UC0331	laurate omega-hydr
25	481	17.6	511 1 B40164	cytochrome P450 4B
26	480	17.6	519 2 T24784	hypothetical prote
27	479.5	17.6	504 2 S50892	cytochrome P450 3A
28	478	17.5	520 2 T24778	hypothetical prote
29	476	17.4	522 2 JG4532	cytochrome P450 4F

30	475	17.4	511 1 B34160	cytochrome P450 4A
31	475	17.4	544 2 T27750	hypothetical prote
32	474.5	17.4	504 1 A32965	cytochrome P450 4A
33	473	17.3	511 1 O4HQB1	cytochrome P450 4B
34	473	17.2	524 2 S29723	cytochrome P450 4F
35	470	17.2	534 2 T23376	hypothetical prote
36	467.5	17.1	520 2 T24777	hypothetical prote
37	466	17.1	515 2 T19576	hypothetical prote
38	466	17.1	524 2 UC7594	cytochrome P450 en
39	466	17.1	524 2 UC7598	cytochrome P450 en
40	465.5	17.1	520 2 T24780	hypothetical prote
41	464	17.0	511 2 S66472	cytochrome P450 4B
42	463.5	17.0	508 1 A36304	cytochrome P450 4A
43	463.5	17.0	510 1 A34160	laurate omega-hydr
44	463	17.0	507 1 A32966	cytochrome P450 4A
45	463	17.0	510 2 T19577	hypothetical prote

ALIGNMENTS

RESULT 1

H84663
probable cytochrome P450 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_reviseion 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84663
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Yencken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, O.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84663
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-520 <STD>
A:Cross-References: GB:AEO02093; NID:92760837; PIDN:AA95305.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g26710
A:Map position: 2

Query Match	Score	DB 2	Length	520
Best local Similarity	99.8%	Pred. No. 2.3e-197		
Matches	519	Conservative	1	Indels
			0	Gaps
			0	
QY	1	MEESSSMFIPKXVLVLSVILSVIVKMSLLMRRPKIEHBFKQIRGPPYHFGNVK	60	
DB	1	MEESSSMFIPKXVLVLSVILSVIVKMSLLMRRPKIEHBFKQIRGPPYHFGNVK	60	
QY	61	ELVGMMLKASHPMPFSEHNLPRVLSFYHMKKIYGATPLVWGGPFRLLTVADPDILREI	120	
DB	61	ELVGMMLKASHPMPFSEHNLPRVLSFYHMKKIYGATPLVWGGPFRLLTVADPDILREI	120	
QY	121	FSKSEFYEKNAHFLVQLGDDGLSLKGEKMAHHRKILSPFPMENLKLLVAVVLSVT	180	
DB	121	FSKSEFYEKNAHFLVQLGDDGLSLKGEKMAHHRKILSPFPMENLKLLVAVVLSVT	180	
QY	181	DMVWMSDKLSENGEVEVDYEMFQILTEDVISRTAFSGSYEDGRAVFRLOAQOMLLCAE	240	
DB	181	DMVWMSDKLSENGEVEVDYEMFQILTEDVISRTAFSGSYEDGRAVFRLOAQOMLLCAE	240	
QY	241	AFOCVFIPGYRFFPFRGNLRSKLDKEIRKSLKLIERRQNAIDEGECKEPAKDLL	300	
DB	241	AFOCVFIPGYRFFPFRGNLRSKLDKEIRKSLKLIERRQNAIDEGECKEPAKDLL	300	
QY	301	GLMTQAKNVTVQDIVECKSFFPAGKQTSNLLTWTTLLSNPBNQAKRDEVAVCGS	360	
DB	301	GLMTQAKNVTVQDIVECKSFFPAGKQTSNLLTWTTLLSNPBNQAKRDEVAVCGS	360	
QY	361	RDVPTKHVYVTKLMSILNLSLTPVATIRBASDVKLGKYPGCTELLPIPIAV	420	
DB	361	RDVPTKHVYVTKLMSILNLSLTPVATIRBASDVKLGKYPGCTELLPIPIAV	420	

QY 421 HHDOAIWGNVNEENPAPAFADGVPAKHPVGFIPGLGVRTICIGNLAILQAKTLAVM 480
 DB 421 HHDOAIWGNVNEENPAPAFADGVPAKHPVGFIPGLGVRTICIGNLAILQAKTLAVM 480
 QY 481 IORFTFHLPYQHAPVLMILLYPQHGAFTIRLTHNHED 520
 DB 481 IORFTFHLPYQHAPVLMILLYPQHGAFTIRLTHNHED 520

RESULT 2

cytochrome P450 - Madagascar periwinkle
 C/Species: Catharantus roseus (Madagascar periwinkle)
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Feb-2001
 C/Accession: T09999
 R/Mangold, U.; Eichel, J.; Batschauer, A.; Lanz, T.; Kaiser, T.; Spangenberg, G.; Werck-
 plant Sci. 96, 129-136, 1994
 A/Title: Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from Catharant
 A/Reference number: Z16915
 A/Accession: T09999
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-524 <MAN>
 A/Cross-references: EMBL:L19074; NID:G404687; PID:G404688
 A/Experimental source: cv. cp3
 C/Genetics:
 A/Gene: CYP72B
 A/Introns: 96/1; 170/3; 252/2; 381/3
 C/Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology
 C/Keywords: heme; iron; metalloprotein
 F/329-492/Domain: cytochrome P450 homology <P45>
 F/470/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 40.8%; Score 113; DB 2; Length 524;
 Best Local Similarity 41.6%; Pred. No. 4.7e-76;
 Matches 216; Conservative 113; Mismatches 164; Indels 26; Gaps 8;
 QY 15 VLSVLSIVYKMSL---WMPRKIEEHFSKQIRGPPYHFFIGNVKELVGMKAS 71
 DB 12 IATITFALVAMAMRVLDMAWFTPKRIKLRQGRGNRYRLVGDVSGKMGQALS 71
 QY 72 HMPFSHNILPRVLSPYHMKIYGATFLVWFGPTFLTVADPDILREISKEFEYKN- 130
 DB 72 KMEFNDIVPRLMFHINTINTYGNSTWGRIPRLHWEPELKEVLTHSSKYQNF 131
 QY 131 EAH-PLVQLBGDGLSLKGEKMAHHRKISPTFMENLKLVPVVLKSTVDWYDKMSDK 189
 DB 132 DVANPLVKFL-TGVSGFGAKMSKRRILISPAFTLEKLSMPLPAFALCYHDMLTWEKI 190
 QY 190 LSENGEVVDVYEWFOILTEDYISRTAFGSSYEDGRAVFLQAOQMLCAEAFQKPIPG 249
 DB 191 AEKQSHHVDIFPEFDVLTSDVSKVAFGSTYDEGKIFRLKEIMDLTIDCKRDVYIPG 250
 QY 250 YRFPFRGNLKSRLDKRIKSLKLEERRQNALDGESECKEPAKDLGLMIQAK-- 307
 DB 251 WSLPTKKNKMKKEINKEI-TDMLRFTINKRKAL-----KAGEGEBDDLGLVLESNIQ 304
 QY 308 -----NTVODIVERCKSPFPAQKQTSNLLTWTTILISMPEWQAKADEYLK 356
 DB 305 EIKQGNKDGMSINDVIEBGLFYFAGQETTVLTLTWTTILSKHEWQERAREEVLQ 364
 QY 357 VCGSRDVPYKQHVYKLTLSMLINESLRIPYIATIRAKSDVKGKYKPCGTTELLIP 416
 DB 365 AFG-KKKEPEFRLNMLKTVSMILYEVRLVPPVIDLTKIHEDTLGSPYTIIPAGTQWMLP 423
 QY 417 IIAVHDOAIWGNVNEENPAPAFADGVPAKHPVGFIPGLGVRTICIGNLAILQAKT 476
 DB 424 TWMLREKSIWGEDAMENPAPAFADGVANATKNNVTYLFPSGPRVLCIGQFALLQAKLG 483
 QY 477 LAVMIORFTFHLPYQHAPVLMILLYPQHGAFTIRRL 515
 DB 484 LAMILORFTFDVAPSYVHAFFITLTVQPFQSGSHVYKKL 522

RESULT 3

cytochrome P450 (CYP72C) - Madagascar periwinkle (fragment)
 C/Species: Catharantus roseus (Madagascar periwinkle)
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Feb-2001
 C/Accession: T10000
 R/Mangold, U.; Eichel, J.; Batschauer, A.; Lanz, T.; Kaiser, T.; Spangenberg, G.; Werck-
 plant Sci. 96, 129-136, 1994
 A/Title: Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from Catharar
 A/Reference number: Z16915
 A/Accession: T10000
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-516 <MAN>
 A/Cross-references: EMBL:L19075; NID:G404689; PID:G404690
 A/Experimental source: cv. cp3
 C/Genetics:
 A/Gene: CYP72C
 C/Superfamily: unassigned cytochrome P450; cytochrome P450 homology
 C/Keywords: heme; iron; metalloprotein
 F/318-481/Domain: cytochrome P450 homology <P45>
 F/459/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 40.7%; Score 110; DB 2; Length 516;
 Best Local Similarity 41.8%; Pred. No. 7.8e-76;
 Matches 217; Conservative 112; Mismatches 164; Indels 26; Gaps 8;
 QY 15 VLSVLSIVYKMSL---WMPRKIEEHFSKQIRGPPYHFFIGNVKELVGMKAS 71
 DB 1 IATITFALVAMAMRVLDMAWFTPKRIKLRQGRGNRYRLVGDVSGKMGQALS 60
 QY 72 HMPFSHNILPRVLSPYHMKIYGATFLVWFGPTFLTVADPDILREISKEFEYKN- 130
 DB 61 NPMEDNDIVPRLMFHINTIKTYGRNSFTWGRIPRLHWEPELKEVLTHSSKYQNF 120
 QY 131 EAH-PLVQLBGDGLSLKGEKMAHHRKISPTFMENLKLVPVVLKSTVDWYDKMSDK 189
 DB 121 DVANPLVKFL-TGVSGFGAKMSKRRILISPAFTLEKLSMPLPAFALCYHDMLTWEKI 179
 QY 190 LSENGEVVDVYEWFOILTEDYISRTAFGSSYEDGRAVFLQAOQMLCAEAFQKPIPG 249
 DB 180 AEKQSHHVDIFPEFDVLTSDVSKVAFGSTYDEGKIFRLKEIMDLTIDCKRDVYIPG 239
 QY 250 YRFPFRGNLKSRLDKRIKSLKLEERRQNALDGESECKEPAKDLGLMIQAK-- 307
 DB 240 WSLPTKKNKMKKEINKEI-TDMLRFTINKRKAL-----KAGEGEBDDLGLVLESNIQ 293
 QY 308 -----NTVODIVERCKSPFPAQKQTSNLLTWTTILISMPEWQAKADEYLK 356
 DB 294 EIKQGNKDGMSINDVIEBCKLFYFAGQETTVLTLTWTTILSKHEWQERAREEVLQ 353
 QY 357 VCGSRDVPYKQHVYKLTLSMLINESLRIPYIATIRAKSDVKGKYKPCGTTELLIP 416
 DB 354 AFG-KKKEPEFRLNMLKTVSMILYEVRLVPPVIDLTKIHEDTLGSPYTIIPAGTQWMLP 412
 QY 417 IIAVHDOAIWGNVNEENPAPAFADGVPAKHPVGFIPGLGVRTICIGNLAILQAKT 476
 DB 413 TWMLREKSIWGEDAMENPAPAFADGVANATKNNVTYLFPSGPRVLCIGQFALLQAKLG 472
 QY 477 LAVMIORFTFHLPYQHAPVLMILLYPQHGAFTIRRL 515
 DB 473 LAMILORFTFDVAPSYVHAFFITLTVQPFQSGSHVYKKL 511

RESULT 4

T09944
 Probable cytochrome P450 protein - Madagascar periwinkle
 N/Alternate names: CYP72 protein
 C/Species: Catharantus roseus (Madagascar periwinkle)
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Feb-2001
 C/Accession: T09944

C:Genetics:
A:Gene: AIC2g46960; F14M4.21
A:Map position: 2
A:introns: 93/1; 168/3; 254/2; 375/3
C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:352-493/Domain: cytochrome P450 homology <P45>
F:471/Binding site: heme iron (Gys) (axial ligand) #status predicted

Query Match	36.1%;	Score 985.5;	DB 2;	Length 526;
Best Local Similarity	37.2%;	Pred. No. 2e-66;		
Matches 196;	Conservative 179;	Mismatches 179;	Indels 35;	Gaps 9;

```

QY 13 VLV,SVTL5LVYKGM5LIMWPKRIEHPKESQOGRGPPHFIQWVKELVGMMLKASSH 72
Db 9 LIVLILILIGIRIFKAFMILWHPFVLVTRLRKNOIGSGNVIYIFGNLSLTKM--KRESH 66
QY 73 PM---PFSNHLIPRVLSFYFHMRRKICYGATFLVWEGPTRLVADPDILREIF5-KSHPYE 128
Db 67 LSLDPSNDIPRILPHYQKMSQYGEFLYWNTEPRICISDPELAKTMLSNKLGPFY 126
QY 129 KNEAHP-LYKOLEGDL5LSLKEGKKAHRKII5PFHEHMKLLVPVVLKSVYDMWDKM- 186
Db 127 K5KAREBAAYLVGSKCLVPEIGADWVRHRRILINPF5IDRLKITTWVOTLKMBEWM 186
QY 187 ---SDLSNGEVEVVDVWEPOLITVEDYISRTAFSS5YEDRAVFRLOAOOMLCABAFO 243
Db 187 KESTRETEHPKIKEMNEEPORLADIIAASAFESS5VEBIEVFRSQMELKRCYTTSLN 246
QY 244 KVPFPGYRPPFPREGNIKSKRLDKELRKSLLKILERRRONAIDGEGECKECPAKDILGLM 303
Db 247 QV5IPEYQYLPFP5NIRVWLKERKDN5IKR15SRLO5-----KSDYGDLLGIL 297
QY 304 IOAKN-----VYODIVEECKSFFPAQKOT5SLTWITLILSMHPWQAKARDEVLR 356
Db 298 LKAYNTEGKERKMSIEIITHECRTFFPGHEHT5SLTAMTLM5LSHQMDQEKLRBIEFR 357
QY 357 VCGSRDVPTRDHYVKKL-----TISM5NE5LFLPYPIATIRPAK5DYKGGYKIPC 409
Db 358 ECKEKEKTP5E5TF5SKIKRKYVFLDMQNWIVESIRLXGVBALARE5VNIKGDDEIFK 417
QY 410 GTELLIPITAVHNDQAIWGDVNEFNPAPAFADGVRBAKHGVGFIPEGLGVRITCIGNTLA 469
Db 418 GTTVAVIPLMKHSDKTLMSDADKKNPKRFPAGV5BRANHHNMLALASVGRACIQGNFV 477
QY 470 TLQAKTLAVMIQRTF5-HLAPTYQAHAPVTLMLLYPQGA5DITRRIL 515
Db 478 M5IAKVLTMILQRFRRF5ISLDBEYKHTPDVANTTIPQYGLGVMLOQPL 524

```

RESULT 9
T05876
cytochrome P450 homolog T29A15.200 - Arabidopsis thaliana
C|Species: Arabidopsis thaliana (mouse-ear cress)
C|Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 28-Jul-2000
C|Accession: T05876
R|Beyan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ho
submitted to the Protein Sequence Database, March 1999
A|Reference number: Z15455
A|Accession: T05876
A|Molecule type: DNA
A|Residues: 1-518 <BEV>
A|Cross-references: EMBL:AL035602; GSPDB:GN00062; ATSP:T29A15.200
A|Experimental source: cultivar Columbia; EAC clone T29A15
C|Genetics:
A|Gene: ATSP:T29A15.200
A|Map position: 4
A|Introns: 95/1; 169/3; 252/2; 376/3
C|Superfamily: human cytochrome P450 CYPAB1; cytochrome P450 homology
C|Keywords: heme; iron; metalloprotein
F|323-487|Domain: cytochrome P450 homology <P45>
F|465|Binding site: heme iron (CVs) (axial ligand) #status predicted

Query Match	34.5%	Score 943;	DB 2;	Length 518;
Best Local Similarity	37.5%	Pred. No. 3.1e-63;		
Matches 196;	Conservative 107;	Mismatches 195;	Indels 24;	Gaps 8;

```

QY 10 TKRVLAVSLVSLVYV----KQMSLMMRPXKIEEHSQGRGPRPHFPGNKVELGM 65
Db 4 IFTIMLTLVLLLEFVSKIMQACWITLLIRPLMLSKRFKQOISGPKYIILGNLSIKM 63
QY 66 MKASSHDM-PFSHUIPLVLSFYHNMRKIKYGATPLVMEGPTFRVLADPDLLIREIF-SK 123
Db 64 KKEADLCVLDPSNDSIDPRVPRQYHQMSQYGDTPFLFTGTPTIYISNHELAKOVLSK 123
QY 124 SEPEYKNAHPLYKOLEBDGILSKSGKNAHKKIISPFHNENIKLVLNPAVLSVDMV 183
Db 124 FGGTII PVKRPVFLFLFKGSLSPFQSDGM-PHRLINLPASMDRLKAMTOPGDCTLRI F 183
QY 184 DKMSDKLENGEV---EVDVWEFQILLEDVISTRTAGSSSEYEDGRVFLQAOOMLCAEA 241
Db 184 EEMR-KQSRNBEVLKIEISKEFKLADILATAGSSYABGIELCSQTELEKYTSS 242
QY 242 FQKVFIPGVRFPPTRGUKSKFKDKERKSLKILERRQNALDGESECKEPA-ANDL 300
Db 243 LNVNFIPIGQYLPPTMLKMLKMKKQMSIKKIDISRLK-----CKKYGYGDDL 295
QY 301 GMLTQ-----KVVTVODIVECKSFPPAGQOTSNLITWTILLSHNPEWQAKADE 353
Db 296 GMLTNAKSNVEYERKGRDBIECKNFYAAAGGTTISILLTWTMLLSHQWQSKREE 355
QY 354 VTRVCGSRDVPKOHVVKLTLSILNESRLYPRPVATITRAKSDVKGKYPCCGRL 413
Db 356 VNEGCGKXKIPBTDIFSLLKLMNVLMESVRLGVPYIKISRATODMKVGHLEIPKGSII 415
QY 414 LRPITAVHDDQIWMNDVNEENPARFADGVPRAAKPVGFIPEFGCVNTEIGONLAIQ 473
Db 416 IIPLKMRDRKAWGEDAEQPNLRPFNGISQATIHNNMLLPFSIGPRACIAKNPAMEVA 475
QY 474 KLTAVMIQRTFHLAPYQHAPVTLMLLYPQHGABITTRRL 515
Db 476 KTVLTMIIQOQFOLSLSPRYKTPVDHDFLEPQGLRVLMLAPL 517

```

```

RESULT 10
D86306
Similar to Cytochrome P450 [imported] - Arabidopsis thaliana
C|Species: Arabidopsis thaliana (mouse-ear cress)
C|Date: 02-Mar-2001 #sequence_rev10n 02-Mar-2001 #text_change 18-Nov-2002
C|Accession: D86306
R|Orthologs: A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, P.
ansen, N.F.; Hughes, B.; Huizar, L.
Native 408; 816-820, 2000
A|Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A|Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallor
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A|Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A|Reference number: A86141; WUID:21016719; PMID:11130712
A|Accession: D86306
A|Status: Preliminary
A|Molecule type: DNA
A|Residues: 1-476 <SNO>
A|Cross-references: GB:AE005172; NID:ig5734759; P|ID:RAD50024.1; GSPDB:GN00141
C|Genetics:
A|Map position: 1
C|Superfamily: Bacillus halodurans cytochrome P450 BH0579; cytochrome P450 homology
C|Keywords: heme; iron; metalloprotein
F|42/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 33.0%; Score 900; DB 2; Length 476;
Best Local Similarity 35.2%; Fred. No. 4,7e-60;
Matches 180; Conservative 107; Mismatches 177; Indels 48; Gaps 5;

```

QY 10 IPKVLVLSVILSV--IVKMSLLMWRPKIEHFSKQIRGPPYHFIQNVKLVGMML 67
 DB 7 VRKVLFIQFLLILNMVRAVNMVWLRPKLEKTKKQSGNSYRILMDQKESNQMDQ 66
 QY 68 KASSHPMPFSNNILPRVLSFYHNRKITYGATFLVWGPPTFLTVADPDLIREIFSSEY 127
 DB 67 VAHSILPLDADPFRMPFLHHVTLVKGKCCFTWYPRVNVMPDRLREIIMSCHLP 126
 QY 128 EK----NEAHPVVKOLEGDGLSLKGEKMAHRIKIIPTFPMENLKLVVVLKSYTDVY 183
 DB 127 PPKPKGSHNHVFL-----SGLLNHEGPKMSKHSILNPARIDKSLILPAFNSCKEML 181
 QY 184 DKMSDKLSNGEVEVDVYEMFOILTEDVISTPAAGSSYEDGRVAVFLQAOQMLCAEAQ 243
 DB 182 EEMERLASAKGTMEIDSWTHCHDLRNLMLASFGSDYKQGIKIFELIQEQLDGLLAR 241
 QY 244 KYEIPGYRPFPPTRGNLKRKLDKEIRKSLKTLIRRRONAIIDGEECKEPPAKDILGLM 303
 DB 242 AVYIGSKFLPFRKFRRLRETRBRKRAFKNMITEKSEIKRGR----- 285
 QY 304 IQAKNVTVODIVECKSFPPAGKQTSNLLTWITLLSMPEWQAKARDEVLRVCSRDV 363
 DB 286 -----AGQVTSLSLFWITLVALSOHQDMQNKARDEISOAFNNR- 324
 QY 364 PPKDHYVVKLTLSMLINESLRLYPPIVATIRAKSDYKLGSKYKIPGTELLIPLVHHH 423
 DB 325 PFEBSHDKVVTMLLHEVRLYSPAYTCITQOEYKLEBFSLPBGVAVTITMLLVHHD 384
 QY 424 QAINQDVNEFNPAFADGVPPRAKHPVGFIPFGIGVTRTCIGQNLAILQAKTLAVMIQR 483
 DB 385 SOLWDDVYKFEKPEFANGVAGATKGRSLFLPSSGPRTCIGQFSLQAKTLAVMLQK 444
 QY 484 FTFHLAPTYQHAPTYMLLIPQHGAPITFRRL 515
 DB 445 FSEVLSPSYTHAPFPAAFTFPQHGAILIIRKL 476

RESULT 11

A96695
 hypothetical protein P5A8.3 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 20-Apr-2001
 C/Accession: A96695
 R/Theologos, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chinn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; anen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A/authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maltz, R.; Marzall, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712
 A/Accession: A96695
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-512 <STO>
 C/cross-references: GB:A005173; NID:g4204278; PIDN:AAD10659.1; GSPDB:GN00141
 C/Genes: P5A8.3
 A/Map position: 1
 A/superfamily: human cytochrome P450 CYP4A1, cytochrome P450 homology
 C/keywords: heme; iron; metalloprotein
 F/58/Binding site: heme iron (axial ligand) #status predicted

Query Match 32.5%; Score 867; DB 2; Length 512;
 Best Local Similarity 35.6%; Pred No. Se-59;
 Matches 185; Conservative 109; Mismatches 189; Indels 36; Gaps 8;
 QY 13 VLVSVILSVIKGMSLLMWRPKIEHFSKQIRGPPYHFIQNVKLVGMML-KASS 71
 DB 12 VLVSVILSVIKGMSLLMWRPKIEHFSKQIRGPPYHFIQNVKLVGMML-KASS 70

QY 72 HPMPSNNILPRVLSFYHNRKITYGATFLVWGPPTFLTVADPDLIREIFSSEYKNE 131
 DB 71 DCSIHNIYVIRLLPHVSVMSKQYGRFIMNGTEPRLCILETEMIKELLTGH----- 123
 QY 132 AHPL-----VKQLEDGLSLKGEKMAHRIKIIPTFPMENLKLVVVLKSYTD 181
 DB 124 -NPVTGSKWLQOQGTGFIIRGLLMANGEAMHQRIMAAPAFRDLKGYAKHVECTM 182
 QY 182 MVDKMSDKLSNGEVEVDVYEMFOILTEDVISTPAAGSSYEDGRVAVFLQAOQMLCAEA 241
 DB 183 MAERLRKEVG-----EVEIEENRRLADLISNTEGSSGDKCKELFSLTVLQRLCAQA 238
 QY 242 FQKFTIIGYRPFPPTRGNLKRKLDKEIRKSLKTLIRRRONAIIDGEECKEPPAKDILG 301
 DB 239 TRHLCPGSRFLPSKVRREIKSLTEVERLIMEIISRKDSVEIGRSSYGD---DLIG 294
 QY 302 LMI-----QAKNVTVODIVECKSFPPAGKQTSNLLTWITLLSMPEWQAKARDEVLR 356
 DB 295 LILNQMNSNNKNNLVQVIMDECKTFFGTGHTSLTLTLMILANFTMODNVREVRQ 354
 QY 357 VCGSRDVPYTDHYVVKLTLSMLINESLRLYPPIVATIRAKSDYKLGSKYKIPGTELLI 416
 DB 355 VCGQDGVPSVEQLSLSLTKVINESLRLYPATLLPRAAFEDIKLGDLIIPKGLSTWIP 414
 QY 417 IIAVHDOAIWQDVNEFNPAFADGVPPRAKHPVGFIPFGIGVTRTCIGQNLAILQAKLT 476
 DB 415 VLAIHNSNELMGEIANEFNBERFTTSFASRR---FMPPAOPRCICQTFAMMEAKII 471
 QY 477 LAVMIQRTFHLAPTYQHAPTYMLLIPQHGAPITFRRL 515
 DB 472 LAMLSKSFSAISENYHAPIVLTIKPYGVQLVKPL 510

RESULT 12

OAKTIO
 laurate omega-hydroxylase (EC 1.14.15.3) cytochrome P450 4A1 - rat
 N/Alternate names: cytochrome P450 1A-omega; cytochrome P452
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 03-Mar-2000
 C/Accession: S01336; A26137; B32965; S21711
 R/Barnshaw, D.; Dale, J.W.; Goldfarb, P.S.; Gibson, G.G.
 FEBS Lett. 236, 357-361, 1988
 A/Title: Differential splicing in the 3'-non-coding region of rat cytochrome P-452 (P4
 A/Reference number: S01336; MUID:88312998; PMID:3410047
 A/Accession: S01336
 A/Molecule type: mRNA
 A/Residues: 1-509 <HAR>
 A/cross-references: EMBL:X07259; NID:g56046; PIDN:CAA30245.1; PID:g56047
 R/Harwick, J.F.; Song, B.U.; Huberman, B.; Gonzalez, F.J.
 J. Biol. Chem. 262, 801-810, 1987
 A/Title: Isolation, complementary DNA sequence, and regulation of rat hepatic lauric
 A/Reference number: A26137; MUID:87109183; PMID:3027069
 A/Accession: A26137
 A/Molecule type: mRNA
 A/Residues: 1-509 <HAR>
 A/cross-references: GB:M4972; NID:g203865; PIDN:AAA41061.1; PID:g203866
 R/Kimura, S.; Hanloka, N.; Matsumaga, E.; Gonzalez, F.J.
 DNA 9, 503-516, 1989
 A/Title: The rat clofibrate-inducible CYP4A gene subfamily I. Complete intron and exo
 am element.
 A/Reference number: A32965; MUID:89356271; PMID:2766932
 A/Accession: B32965
 A/Molecule type: DNA
 A/Residues: 1-340, E, 342-509 <KIN>
 A/cross-references: EMBL:M57718; NID:g203786; PIDN:AAA41038.1; PID:g203787
 R/Okita, R.T.; Okita, J.R.
 Arch. Biochem. Biophys. 294, 475-481, 1992
 A/Title: Characterization of a cytochrome P450 from Di-(2-ethylhexyl) phthalate-treat
 A/Reference number: S21711; MUID:92231570; PMID:1567203
 A/Accession: S21711
 A/Status: preliminary
 A/Molecule type: Protein

A:Residues: 1-22 <OK>
 A:Experimental source: strain Sprague-Dawley
 A>Note: this sequence was confirmed by protein sequencing
 C:Genetics:
 A:Gene: CYP4A1
 A:introns: 63/3; 112/1; 127/1; 169/3; 211/2; 263/1; 298/3; 362/2; 407/1; 428/3; 454/2
 C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology
 C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; fatty acid oxidatio
 F:317-478/Domain: cytochrome P450 homology <P45>
 F:456/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 19.2%; Score 523; DB 1; Length 509;
 Best Local Similarity 28.4%; Pred. No. 1.4e-31;
 Matches 151; Conservative 105; Mismatches 218; Indels 58; Gaps 17;

```

QY 5 SSSMFIPKVLVLSVLIVYKMSLLMWRPKLIEHFSKQIGRPYHFTGNV-----59
DB 14 SLSGFLQVASVIGLL--LLYKAVQFYLRQMLK--AFQPPSPFPFHFQHKQFGD 68
QY 60 KELVGMMLKASHMPFSGNITLPRVLSFYHMRKIYGATFLVWF--GPTFRLTVADPDILR 118
DB 69 KELQGITCVENPFSAP-----PR-----WFGSKAVLYVDDPDYMK 105
QY 119 EIFSKSEFEKNEAHPVLRQLEGGDLSLKEKMAHRRKIIISPTFMENIKLIVPVVLS 178
DB 106 VILGRSD--PKANGVYRLAPWIGYGLLLNGQPFQHRRLTPAFHYDLIKPYVKNMADS 164
QY 179 VTDMDVKMSDKLSENGEVVDVYEMFQILTEDVISRTAF---GSSYEDGRAVRLQAQOM 235
DB 165 IRLMDKWEQIAGDSSIE--TFQHSIMLTVDYKCAFSHNSVOVDGNYKSYQAIQIN 222
QY 236 L--LCAEAFQKVFIPGYAEF--FPTRGMLKSR--KLDEIRKSLKLIERRQNALDGE 289
DB 223 LNDLFHSRVNRNIFHQNDITYFSSNGHLFNRACQALADHTDGVYLRKQDLQA--GELE 280
QY 290 ECKEPAKQDLGLMTQAK-----NTVQDIYVECKSFPPAGKQTSNLTWTTLISMP 344
DB 281 KYKKRRRLDFLDILLFARMENGDSLSPKDLAEVDITFMFGHDTTASGVSWIYALATHP 340
QY 345 EMOAKARDEVLRVCGSRDPVTKDHYVVKLTLSMILNESLRTPVATIRRAKSDVKL-G 403
DB 341 KHQQRCEBVSVLGDSSTIWDHDOIPTTMCIKALMLPPVPGVIELSTVTFPD 400
QY 404 GYKIPCGTELLIPITAVHDDQAIWGNVNEPNRPAPDGVPPRAKGVGTFPGGLGYTC 463
DB 401 GRSILPKGIQVTLSTIYGLHNPKVWPNP--EVPDPSPRFAPDSPR--HSHSFLPFGSGARN 456
QY 464 IGQNLAILQAKTLTAVMIOQFTFHLAPYCHAPVLMILYPOHGAPITFRL 515
DB 457 IGKQPMSEMKVYVLTLLRPELLDPPTKVPFIPRLVLSKNGIYLYLKKL 508

```

RESULT 13

S47553

Cytochrome P450 CYP4A - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 04-Mar-2000
 C:Accession: S47553

R:Henderson, C.J.; Bammler, T.; Wolf, C.R.
 Biochim. Biophys. Acta 1200, 182-190, 1994
 A:Title: Dedicated amino acid sequence of a murine cytochrome P-450 Cyp4a protein: develop
 A:Reference number: S47553; PMID:94304927; PMID:8031839
 A:Accession: S47553
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-509 <HEN>
 A:Cross-references: GB:X69296; NID:9312695
 A>Note: this ORF is not annotated in GenBank entry MNCYP4A, release 109.0
 C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology
 C:Keywords: chromoprotein; heme; iron; metalloprotein; transmembrane protein
 F:317-478/Domain: cytochrome P450 homology <P45>
 F:456/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 18.6%; Score 507; DB 2; Length 509;
 Best Local Similarity 28.0%; Pred. No. 2.2e-30;
 Matches 149; Conservative 107; Mismatches 219; Indels 58; Gaps 17;

```

QY 4 ESSGWFIPKVLVLSVLIVYKMSLLMWRPKLIEHFSKQIGRPYHFTGNV---60
DB 13 DLSJGFLQVASVIGLL--LLYKAVQFYLRQMLK--AFQPPSPFPFHFQHKQFGD 67
QY 61 --ELVGMMLKASHMPFSGNITLPRVLSFYHMRKIYGATFLVWF--GPTFRLTVADPDIL 117
DB 68 DQELQEVSCIEPPAPF-----PR-----WFGSKAVLYVDDPDYMK 104
QY 118 REIFSKSEFEKNEAHPVLRQLEGGDLSLKEKMAHRRKIIISPTFMENIKLIVPVVLS 177
DB 105 KVLGRSD--PKANGVYRLAPWIGYGLLLNGQPFQHRRLTPAFHYDLIKPYVKNMAD 163
QY 178 VTDMDVKMSDKLSENGEVVDVYEMFQILTEDVISRTAF---GSSYEDGRAVRLQAQOM 234
DB 164 SIRLMDKM--ERLADQ--DSIEIFQHSIMLTVDYKCAFSHNSVOVDGNYKSYQAIQIN 221
QY 235 ML--LCAEAFQKVFIPGYAEF--FPTRGMLKSR--KLDEIRKSLKLIERRQNALDGE 288
DB 222 DNLNLFHSRVNRNIFHQNDITYKLSNGRLAKQACQALADHTDGVYLRKQDLQA--GELE 279
QY 289 ECKEPAKQDLGLMTQAK-----NTVQDIYVECKSFPPAGKQTSNLTWTTLISMP 343
DB 280 EKIKRRRLDFLDILLFARMENGDSLSPKDLAEVDITFMFGHDTTASGVSWIYALATHP 339
QY 344 EMOAKARDEVLRVCGSRDPVTKDHYVVKLTLSMILNESLRTPVATIRRAKSDVKL-402
DB 340 PDHQQRCEBVSVLGDSSTIWDHDOIPTTMCIKALMLPPVPGVIELSTVTFPD 399
QY 403 GYKIPCGTELLIPITAVHDDQAIWGNVNEPNRPAPDGVPPRAKGVGTFPGGLGYTC 462
DB 400 DGRSLPKGVQVTLSTIYGLHNPKVWPNP--EVPDPSPRFAPDSPR--HSHSFLPFGSGARN 455
QY 463 IGQNLAILQAKTLTAVMIOQFTFHLAPYCHAPVLMILYPOHGAPITFRL 515
DB 456 IGKQPMSEMKVYVLTLLRPELLDPPTKVPFIPRLVLSKNGIYLYLKKL 508

```

RESULT 14

T24783

hypothetical protein T10B9.10 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Feb-2001
 C:Accession: T24783

R:Gardner, A.
 submitted to the EMBL Data Library, March 1995
 A:Reference number: Z19935
 A:Accession: T24783
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-518 <WIL>
 A:Cross-references: EMBL:Z48717; PIDN:CAA88609.1; GSPDB:GN00020; CESP:T10B9.10
 A:Experimental source: clone T10B9
 A:Genes: CESP:T10B9.10
 A:Map position: 2
 A:introns: 62/3; 173/2; 206/3; 392/2; 440/2
 C:Superfamily: human cytochrome P450 CYP2A5; cytochrome P450 homology
 C:Keywords: heme; iron; metalloprotein
 F:323-486/Domain: cytochrome P450 homology <P45>
 F:464/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 18.4%; Score 503; DB 2; Length 518;
 Best Local Similarity 26.6%; Pred. No. 4.6e-30;
 Matches 148; Conservative 109; Mismatches 180; Indels 120; Gaps 20;

```

QY 17 SVLSVLIVYKQ--MSLLMWRPKLIEHFSKQIGRPYHFTGNVKELVGMMLKASHMP 75
DB 4 SILNAILFVGIIISYVW----IWSFWIRKGVKGRGLPFLGVYIHK-----45

```

```

QY 76 FSHNLLPVLSFYHMKRYGATFLVWFSPTRLLTVADPDLIREIFSK--SEFY----- 127
DB 46 FTNENQALKF--SEWTKKIGPVYIGTEGVEKTLVSDPEFVVEVVKQFDFYGRKTLA 104
QY 128 -----EKNEAHPVLVKOLEGGDLSLKGEKMAHHRKTIISPTFMENIKLIVPVVLSYVDM 182
DB 105 IQGDPNKXKRVPLV-----AAQGRMKRLRLTLASPTFSNKSILRKIVGESVTEL 155
QY 183 VDKMSDKLSENGEYVYVYEWFOILTEVDYISRTAFSS-----YEDGRAVPRLOAQ 233
DB 156 V-RSLERKASAEKKT-LDMLEYIQEFTWDIIGKAMGEEKSLMRANMDKVTIIEGEN 213
QY 224 QMLCAEAFQKVFIPGYRFPPTGNLK-SRLDKERKSLKLIERRON--AIDGEGE 289
DB 214 NVFMTSGIFPFGVIALRNIFAKFPSLQMATDIOIIEKALNKRLERDEKAGIEPSSGE 273
QY 290 ECKEPAKDILGLMIQA-----KNVTQDIVECKSPFFAG 325
DB 274 -----PODFIDLFLDASTYDFEGEABODFAKSEVLYKVDKHLTDELIIGQLFVLLAG 327
QY 326 KOTTSNLLTWTTILSNHPEWQAKARDEVLRVCGSRDVPFKDHVVKLKTLSMILNESLRL 385
DB 328 YDTPLSLSYSSYLATHPETIQKLOGEVDRECPDEV--TFDQISKLKYLECVKXKALRL 386
QY 386 YPPIVAT--RRAKSDVKLGKYKIPCGTELLIPTIAVHDOATMGDVFNEFPAR----- 438
DB 387 YP--LASLVNHRKCLKTNNVLMGEIENGTINVDWLSLHDPKVMGDVNEFPKERMESG 444
QY 439 ----FADGVPRAAKHPVGFIPFGIGVTRCIGNLAILQAKLTIAVNIQRTFLHAF----- 490
DB 445 DELFPKAG-----GYLPGMGPRICIGMLAMEMGMLLTNILKVTFTTETETVI 495
QY 491 ----TYOHAPTVLM 501
DB 496 PLKLVGTATIASSVLL 512

```

RESULT 15

```

T24779
hypothetical protein T10B9.3 - Caenorhabditis elegans
C1:Species: Caenorhabditis elegans
C2:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Feb-2001
C3:Accession: T24779
R3:Gardner, A.
submitted to the EMBL Data Library, March 1995
A1:Reference number: Z19935
A2:Accession: T24779
A3:Status: preliminary; translated from GB/EMBL/DBJ
A4:Molecule type: DNA
A5:Residues: 1-518 <M1>
A6:Cross-references: EMBL:Z48717; PIDN:CAA8605.1; GSPDB:GN00020; CESP:T10B9.3
A7:Experimental source: clone T10B9
C1:Genetics:
A1:Gene: CESP:T10B9.3
A2:Map position: 2
A3:Introns: 60/3; 171/2; 204/3; 390/2; 438/2
C4:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C5:Keywords: heme; iron; metalloprotein
F1:21-485/Domain: cytochrome P450 homology <P45>
F2:463/Binding site: heme iron (Cys) (axial ligand) #statue predicted

```

Query Match

```

Best Local Similarity 18.1%; Score 495; DB 2; Length 518;
Matches 142; Conservative 101; Mismatches 197; Indels 92; Gaps 17;

```

```

QY 15 VLSVILSVIVKMSLIMRPRKIEHPSKQIGRPYHFFIGNVKELVGMMLKASSHPM 74
DB 1 MIFVLISAVLLGVFTYGV-----LMSYIRKIKIGP--RGFG-----IGMLIQTLIDHEN 49
QY 75 PFSHNLLPVLSFYHMKRYGATFLVWFSPTRLLTVADPDLIREIFSK--SEFYEKNEA 132
DB 50 P-----PFLKYRDMTKQYGVYGFTEGPOQTWIISEPEWNEIFKKQFQDNFYGR-KL 100

```

```

QY 133 HPLVKOLEGD---GLISLKGEKMAHHRKTIISPTFMENIKLIVPVVLSYVDMVDMKSD 189
DB 101 RPIIGDPEDKRVNIFSTQGRKMKRLRLTLSSPSFNNISLRKVRNSVQSGTEILWNIOK 160
QY 190 LSENGEYVYVYEWFOILTEVDYISRTAFSSYED-----GRAVPRLOAQML----- 237
DB 161 VRKNEDIDMLIV--YQETLGVISRIALQSSNNFKNPILPKVQAIFNGSHVFLINGI 218
QY 238 ---CAEAFQV-----FIPGYRFPPTGNLKSRLDKERKSLKLIERRONAIIDGE 287
DB 219 FPLIAGVFRKMSKMLPASFIAPKIFDL-----IEVAQARIDQAKDEI--K 264
QY 288 GECKEPAKDILGLMIQA-----KNVTQDIVECKSPFFAG 322
DB 265 GVEPGEF--ODFIDLFLDARVPDVKLLSEANEDFAKSSVVKINKELTDELIAGQFVFL 322
QY 323 FAGKOTTSNLLTWTTILSNHPEWQAKARDEVLRVCGSRDVPFKDHVVKLKTLSMILNES 382
DB 323 AAGPDTLALSLSATYLLATHPEIQKLOGEVDRECPDEI--FPDHLSTKLKYLECVKXKALRL 381
QY 383 LRLYP--PIVATTIRAKSDVKLGKYKIPCGTELLIPTIAVHDOATMGDVFNEFPARFAD 441
DB 382 LRLYPDTTANTRKCMRETTINGVNFDEGNITQVDWTLHNPRIENGVEVDPEKERMEN 441
QY 442 GVPRAAKHPVGFIPFGIGVTRCIGNLAILQAKLTIAVNIQRTFLHAF----- 493
DB 442 GACCHLEHNSYIFPFGSGPRQCIQMLAQMEQKILIAQLKKEYSPRTTQVQ 493

```

Search completed: May 19, 2004, 10:27:52
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 10:19:50 ; Search time 17 Seconds
(without alignments)

1592.734 Million cell updates/sec

Title: US-09-992-901-2

Perfect score: 2730
Sequence: 1 NEBSSSWFPMVLSVIL.....LTPQHGAPITFRLLTHED 520

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1101	40.3	524	1 CP72 CATRO	Q05047 catharanthu
2	523	19.2	509	1 CP41 RAT	P08516 rattus norv
3	503	18.4	518	1 YRVA CAEEL	Q27519 caenorhabdi
4	495	18.1	518	1 YRVA CAEEL	Q27515 caenorhabdi
5	492	18.0	511	1 CP45 RABIT	P14579 oryctolagus
6	487.5	17.9	504	1 CP38 MOUSE	Q64459 mus musculu
7	487.5	17.9	506	1 CP44 RABIT	P16611 oryctolagus
8	482	17.7	537	1 CP86 RAT	P51871 rattus norv
9	481.5	17.6	504	1 CP41 MOUSE	Q93ma7 mus musculu
10	481.5	17.6	519	1 CP4Y HUMAN	Q02928 homo sapien
11	481	17.6	511	1 CP4B RAT	P15129 rattus norv
12	480	17.6	519	1 YRVA CAEEL	Q27520 caenorhabdi
13	479.5	17.5	520	1 YRVA CAEEL	Q64481 mus musculu
14	478	17.5	520	1 YRVA CAEEL	Q27514 caenorhabdi
15	477	17.5	511	1 CP4B HUMAN	P13584 homo sapien
16	477	17.5	524	1 YS24 CAEEL	Q09653 caenorhabdi
17	476.5	17.5	524	1 CPFB HUMAN	Q9hb16 homo sapien
18	476	17.4	522	1 CPFB RAT	P51869 rattus norv
19	475.5	17.4	510	1 CP46 RABIT	P14580 oryctolagus
20	474.5	17.4	504	1 CP43 RAT	P20816 rattus norv
21	473	17.3	511	1 CP4T RABIT	P14581 oryctolagus
22	470	17.2	524	1 YRVA CAEEL	P33774 rattus norv
23	467.5	17.1	520	1 YRVA CAEEL	Q27513 caenorhabdi
24	466	17.1	524	1 YRVA CAEEL	Q9h682 homo sapien
25	465.5	17.1	520	1 YRVA CAEEL	Q27517 caenorhabdi
26	464	17.0	520	1 CP4B MOUSE	Q64462 mus musculu
27	464	17.0	520	1 CP4B HUMAN	P98187 homo sapien
28	463.5	17.0	508	1 CP4B RAT	P24464 rattus norv
29	463	17.0	507	1 CP43 RAT	P20817 rattus norv
30	459	16.8	478	1 CP43 DROME	Q9y470 drosophila
31	458	16.8	526	1 CPFB RAT	P51870 rattus norv
32	455	16.7	503	1 CP39 RAT	P51338 rattus norv
33	454.5	16.6	504	1 CP32 RAT	P05183 rattus norv

ALIGNMENTS

RESULT 1	CP72 CATRO	STANDARD;	PRT;	524 AA.
AC	005047;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cytochrome P450 72A1 (EC 1.3.3.9) (CYPLXIII) (Secologanin synthase)			
DE	(SLS).			
GN	CYP72A1 OR CYP72 OR P450CR3.			
OS	Catharanthus roseus (Roe's periwinkle) (Madagascar periwinkle).			Q70537 mesocricetu
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			Q29496 ovis aries
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;			Q9va27 drosophila
OC	Lamiales; Gentianales; Apocynaceae; Rauvolfioideae; Vincaceae;			Q27518 caenorhabdi
OC	Catharanthus.			P29981 blaberus di
OX	NCBI_TaxID=4058;			P78128 oryctolagus
RN	[1]			P78129 homo sapien
RP	SEQUENCE FROM N.A.			Q9v779 drosophila
RC	STRAIN=cv. CP3A;			P04800 rattus norv
RA	Vetter H.-P., Mangold U., Schroeder G., Marner F.-U.,			Q27589 drosophila
RA	Meier-Reichardt D., Schroeder U.,			Q64464 mus musculu
RT	"Molecular analysis and heterologous expression of an inducible			P47787 sus scrofa
RT	cytochrome P-450 protein from periwinkle (Catharanthus roseus L.).";			
RL	Plant Physiol. 100:998-1007(1992).			
RL	[2]			
RP	SEQUENCE OF 463-524 FROM N.A.			
RC	STRAIN=cv. G. Don;			
RA	MDLLINE=93283641; PubMed=8507838;			
RA	Meijer A.H., Souer E., Verpoorte R., Hoge J.H.C.;			
RT	"Isolation of cytochrome P-450 cDNA clones from the higher plant			
RT	Catharanthus roseus by a PCR strategy.";			
RL	Plant Mol. Biol. 22:379-383(1993).			
RL	[3]			
RP	FUNCTION.			
RC	STRAIN=cv. CP3A;			
RA	MDLLINE=20575722; PubMed=11135113;			
RA	Immler S., Schroeder G., St-Pierre B., Crouch N.P., Hotze M.,			
RT	Schmidt J., Strack D., Matern U., Schroeder J.;			
RT	"Indole alkaloid biosynthesis in Catharanthus roseus: new enzyme			
RT	activities and identification of cytochrome P450 CYP72A1 as			
RL	secologanin synthase.";			
RL	Plant J. 24:797-804(2000).			
RC	-1- FUNCTION: Converts loganin into secologanin.			
CC	-1- CATALYTIC ACTIVITY: Loganin + NADPH + O(2) = secologanin + NADP(+) + 2 H(2)O.			
CC	-1- PATHWAY: Indole alkaloids biosynthesis.			
CC	-1- SUBCELLULAR LOCATION: Vacuolar membrane (Probable).			
CC	-1- TISSUE SPECIFICITY: Upper and lower leaf epidermis.			
CC	-1- SIMILARITY: Belongs to the cytochrome P450 family.			
CC	-1- This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/)			

or send an email to license@lsb-sib.ch.

CC EMBL; L10081; AAA33106.1; -
 DR EMBL; X69775; CAA49430.1; -
 DR PIR; S35168; S35168.
 DR PIR; T09944; T09944.
 DR HSSP; P14779; 1JPZ.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450.1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Transmembrane;
 KW NADP; Heme; Alkaloid metabolism.
 FT TRANSMEM 12 32 POTENTIAL.
 FT METAL 470 470 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT VARIANT 190 190 I -> L.
 FT VARIANT 194 194 Q -> E.
 FT VARIANT 223 223 E -> D.
 FT VARIANT 312 312 K -> R.
 FT VARIANT 318 318 S -> T.
 FT VARIANT 403 403 V -> I.
 FT VARIANT 405 405 K -> E.
 FT VARIANT 411 411 S -> P.
 FT VARIANT 411 411 S -> P.
 SQ SEQUENCE 524 AA; 60557 MW; EFSDB64E43C751B8 CRC64;

Query Match 40.3%; Score 1101; DB 1; Length 524;
 Best Local Similarity 41.4%; Pred. No. 4e-63;
 Matches 215; Conservative 112; Mismatches 166; Indels 26; Gaps 8;

QY 15 VLAVLSIVYKGSLL---WRRPKIEEHSKSGKGRPHYFPGINVKELVOMLKAS 71
 DB 12 IAAITFALVWAMRVDMAMFTPKRLKRLQGGFRGNFRLVGVDSGMMGEALS 71
 QY 72 HPMFPHNILPRVLSFYHWRKIKYATFLVFGFTPLTVADPDLREIFSKSEFYKX 130
 DB 72 KPMENNDIVRLMHNHINTYGRNSFTWNGRIPRIHWEDELKEVLTHSKYQKNF 131
 QY 131 EAH-PLVQLESGDGLSLKSKAKAHNRKISPTFMENLTLVPLVYLKSTVDWVDSK 189
 DB 132 DVNPLVYKFL-TGVSGFEQKMKSKRRRIISPAETLEKLSMPAPAICYHDLTWKRI 190
 QY 190 LSENGEVVDYVEFQITTEDEVISRTAFSGSYEDGRVAFLOQOMLLCAEAFQYFPG 249
 DB 191 AEKQSHVDLFFPTDVLSIVIKVAFSGTYEGSKIFLLKELMDLITDCRDVYIG 250
 QY 250 YRFPFRGNLKRKLDEIKRSLIKLIERRQNAIDEGEGCEKPAKDLGLMIQAK-- 307
 DB 251 WSYLPTKKNKMKKEINKEI-TDMRFLINKRMKAL-----KAGBPGEDDLGLVLSNTQ 304
 QY 308 -----NVTVODIVECKSEFFPAKQTTNLTTLTTLISMHPEWQAKRDEVL 356
 DB 305 ELOKQGNKQDQMSINDVIECKLFYRAGCETTVLLTWTLISKHEPQGEAREVQ 364
 QY 364 AFG-KNKEPERINLKYSMILVEVLRLYPVLDLRIYKQTKGSIYIPAGTQVMP 423
 DB 365 VCGSEVPTKDVVYKLTLSMILNESLRLPYVATIRAKSDYKGGYKIPGTELLP 416
 QY 417 IIAVHDAINGNDVNEFNPARFPDGVPRAKHGVGIPGLQVGTIGONLLIQAOKT 476
 DB 424 TWMLEKRSIWGEDMEENRPFQGVANAKKNVTVLIPFSGRVLGQNPALLQAKG 483
 QY 477 LAVMIQRTFLAPTYQAPVLMILVYPOHGAPITFERL 515
 DB 484 LAMTLQRFKFDVAPSVYAAPFLLTVGQFGSHVYKKL 522

RESULT 2
 CP41_RAT STANDARD; PRT; 509 AA.
 AC P08516;
 DT 01-AUG-1988 (rel. 08, Created)
 DT 01-AUG-1988 (rel. 08, Last sequence update)
 DT 28-FEB-2003 (rel. 41, Last annotation update)

DE Cytochrome P450 4A1 (EC 1.14.15.3) (CYP1A1) (lauric acid omega-
 hydroxylase) (P450-1A-omega 1) (P452).
 GN CYP4A1 OR CYP4A-1.
 OS Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Chordata; Rodentia; Sciurognathi; Muridae; Rattus.
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_Taxid=10116;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=87109183; PubMed=3027069;
 RA Hardwick J.P., Song B.-U., Huberman E., Gonzalez F.J.;
 RT "Isolation, complementary DNA sequence, and regulation of rat hepatic
 RT lauric acid omega-hydroxylase (cytochrome P-450A omega)."
 RT Identification of a new cytochrome P-450 gene family.";
 RT J. Biol. Chem. 262:801-810(1987).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Liver;
 RX MEDLINE=88312998; PubMed=3410047;
 RA Barnshaw D., Dale J.W., Goldfarb P.S., Gibson G.G.;
 RT "Differential splicing in the 3' non-coding region of rat cytochrome
 RT P-452 (P450 IVA1) mRNA.";
 RT FEBS Lett. 236:357-361(1988).
 [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89356271; PubMed=2766932;
 RA Kimura S., Hanloka N., Matsunaga E., Gonzalez F.J.;
 RT "The rat clofibrate-inducible CYP4A gene subfamily. I. Complete
 RT intron and exon sequence of the CYP4A1 and CYP4A2 genes, unique exon
 RT organization, and identification of a conserved 19-bp upstream
 RT element.";
 RT DNA 8:503-516(1989).
 [4]
 RP COVALENT HEME ATTACHMENT.
 RX MEDLINE=21179192; PubMed=11139583;
 RA Hoch U., Ortiz de Montellano P.R.;
 RT "Covalently linked heme in cytochrome P450A fatty acid
 RT hydroxylases.";
 RT J. Biol. Chem. 276:11339-11346(2001).
 [5]
 RP COVALENT HEME ATTACHMENT, AND MUTAGENESIS OF GLU-320.
 RX MEDLINE=21935389; PubMed=11821421;
 RA Lehn U.A., Hoch U., Ortiz de Montellano P.R.;
 RT "Autocatalytic mechanism and consequences of covalent heme attachment
 RT in the cytochrome P450A family.";
 RT J. Biol. Chem. 277:12755-12761(2002).
 CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. In liver microsomes, this enzyme is involved in an
 CC NADPH-dependent electron transport pathway. It oxidizes a variety
 CC of structurally unrelated compounds, including steroids, fatty
 CC acids, and xenobiotics.
 CC -1- CATALYTIC ACTIVITY: Octane + reduced rubredoxin + O(2) = 1-octanol
 CC + oxidized rubredoxin + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -1- INDUCTION: By clofibrate.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).

DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KM Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KM Microsome; Endoplasmic reticulum.
 FT BINDING 320 320 HEME (COVALENT).
 FT METAL 456 456 IRON (HEME AXIAL LIGAND).
 FT MUTAGEN 320 320 E->A: LOSS OF HEME BINDING.
 FT CONFLICT 341 341 K -> E (IN REF. 3).
 SO SEQUENCE 509 AA; 58214 MW; 11E5E102476ADDC2 CRC64;

Query Match 19.2%; Score 523; DB 1; Length 509;
 Best Local Similarity 28.4%; Pred. No. 2,6e-26;
 Matches 151; Conservative 105; Mismatches 218; Indels 58; Gaps 17;

QY SSSWFIKVLVSLVSLVVKGMSLIMWRPKIEHPSKQIRGPPYHFGNVLK-----59
 DB SSGFLGVAVLGLL--LLLVKAVQVFLQKQWLLK--AFQOPSPFFHFWFGHKQFGD 68
 QY 60 KELVGMKLKASSHPMPSSNILLPRVLSFYHMKRTIGATPLVWF--GPTFLVADPDLIR 118
 DB 69 KELQOIMTCVENPSPAF---PR-----WFGSKAYLLVDPDYMK 105
 QY 119 EIFKSEFEYENKEMHPLVKQLEGDGLSLKGEKVAHHRKISPTFEMENIKLIVPVLS 178
 DB 106 VILGRSD-PRANGVYRLAPWIGYGLLLNGQPMQHRKMLTPAFHYDILKRYVNMADS 164
 QY 179 VTDWVDMKSKLSENGEVEVDVYEWQILTEVDISRTAF--GSSYEDGRAVFLQAOQM 235
 DB 165 IRLMLDKMEQLAGQDSSIR--IFQHSIMTLDTVMCAFSHNSVQVDNYSYIOAIGN 222
 QY 236 L--LCAPAFQKVPFPGRF--FPTRGMLKSR--KLQKEIRKSLKIEHRRONADGEGE 289
 DB 223 LNDLFHSRVRNIFQNDNTINFSNGELFNRAQCLADHTDGVYIKRQOLQNA--GELE 280
 QY 290 ECKEPAKDLGLMIOAK-----NVTVDIVECKSFPPAGKOTSNLLTWTTILLSMHP 344
 DB 281 KVKKGRRLDPLDILLARMENGSLSDKDLRAVDFTFMEGHDTTASGVSWIFVALATHP 340
 QY 345 EWQAKADEVLRVCGSDVPTKDVVVKLSTLMSILNESLRIPRYVATIRPAKSDVYKL-G 403
 DB 341 KHQORCREEYQSVIGDSSITWDHLDOIPTYTWCIEKALRLYPPVGIVRELSTSTFEPD 400
 QY 404 GYKIPCGTELLIFPIAVHDOAIWGDVNEFNPNARFADGVPRAAKPPVFIFPGAGVTC 463
 DB 401 GRSIPKGIQVTLSTYGLHHPKWPNP-EVFDSPRFPDPSPR---HSHFELPFGSGARNVC 456
 QY 464 IGMNLAIIQAKLTLAVMIQRTFHLAPTYGHAFVLMLLYPOGADITRRRL 515
 DB 457 IGKQFAMSEMKVIALTLIRFELLDPKVPILPRLVLKSKNGIYLYLKKL 508

RESULT 3

QY RVAA_CAEEL STANDARD; PRT; 518 AA.
 AC Q27519;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative cytochrome P450 CYP13A7 (EC 1.14.-.-).
 GN CYP13A7 OR T10B9.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_Taxid=6239;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol NZ;
 RA Gardner A.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-cholate
 CC monooxygenases. They oxidize a variety of structurally unrelated
 CC compounds, including steroids, fatty acids, and xenobiotics.

CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Z48717; CA88609.1; -.
 DR PIR; T24783; T24783.
 DR HSP; P14779; JUPZ.
 DR Wormsep; T10B9.10; CE01655.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KM Hypothetical protein; Oxidoreductase; Monooxygenase; Heme.
 FT METAL 464 464 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SO SEQUENCE 518 AA; 58999 MW; CC04283EF87B9EA7 CRC64;

Query Match 18.4%; Score 503; DB 1; Length 518;
 Best Local Similarity 26.6%; Pred. No. 5e-25;
 Matches 148; Conservative 109; Mismatches 180; Indels 120; Gaps 20;

QY 17 SVLSLVIVYKQ-MSLIMWRPKIEHPSKQIRGPPYHFGNVLKGVGMKLKASSHPMP 75
 DB 4 SIMLAIAIFGISTYLM-----IMSWIKYKRGKRGPLGVIRK-----45
 QY 76 FSHNILPRVLSFYHMKRTIGATPLVWFGPTPLTVADPDLIRLIFSK--SEFY-----127
 DB 46 FTVYENPQALKF--SEWTKKYPGYGTEGEVETLVISDPEFHEVFGQDFNFKULTA 104
 QY 128 -----EKNEHNPVKQLEGDGLSLKGEKVAHHRKISPTFEMENIKLIVPVLSYTD 182
 DB 105 IQGDPNKNKRVPLV-----AAQGHMKRLRTLASPTFNKSLRKMTGVESVTEL 155
 QY 183 VDKMSDKLSENGEVEVDVYEWQILTEVDISRTAFSS-----YEDGRAVFLQAO 233
 DB 156 V-RSLEKASAEKGT-IDMLEVYQEFMDIIGMAQGEKSLMFRNPMLDKVTIFKEGN 213
 QY 234 QMLICAEAFQKVPFPGRFPTRGMLKSR--KLQKEIRKSLKIEHRRON--AIDGEGE 289
 DB 214 NVFMISGIFPFVGAIRNIFAKPRLQVATDIQSLIERKLRRLQREADEAGLEPSGE 273
 QY 290 ECKEPAKDLGLMIOA-----KNVTVDIVECKSFPPAG 325
 DB 274 -----PQDIDFLQARSVDFPBGSAQDFAKSEVLKVDHLPFDEILIGLFPVLLAG 327
 QY 326 KQTTSNLLTWTTILLSNHPWQAKADEVLRVCGSDVPTKDVVVKLSTLMSILNESLR 385
 DB 328 YDTLASLSYSYILATHEPIQKQOEVDRECPDEV--TFQOLSLKTLKCVVREALRL 386
 QY 386 YPRIVARTI---RRAKSDVVLGKYKIPCGTELLIFPIAVHDOAIWGDVNEFNPNAR 438
 DB 387 YP--LASLVNRRKCLKTNNVGLMEIAGNINIVDTWSLHDKRVQMDVNEKPRRWSG 444
 QY 439 ----FADGVPRAAKHPVGF-PFGAGVATTCIGQNLATLQAKLTLAVMIQRTFHLAP 490
 DB 445 DELFFKAG-----GYLPFGMGPRICIGMLAMEMKMLTNILKNYTFEITPEVTL 495
 QY 491 -----TYQNAFPTVLM 501
 DB 496 PLKLVGATTAAPSSVLL 512

RESULT 4

QY RVV3_CAEEL STANDARD; PRT; 518 AA.
 AC Q27515;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

SQ SEQUENCE 511 AA; 58357 MW; 11D174BFC8BA268 CRC64;
 Query Match 18.0%; Score 492; DB 1; Length 511;
 Best Local Similarity 27.7%; Pred. No. 2.5e-24;
 Matches 146; Conservative 103; Mismatches 219; Indels 60; Gaps 17;

QY 16 LSVLSLVIVKMSL-----WMPRKIEHFSSQSGRGPYHFGNVELTG 64
 DB 15 LSGLLQVAAALGLLILKAQOLYRQMLRALQO-----PCCPFFMLHSHS--- 65
 QY 65 MMLKASHPMPESHN-ILPVLSPFYHMKRYGATFLVWGPT-FRLTVADPDLREIFS 122
 DB 66 -----FQNNQQLQOITK-----WKEKFPACPHWIGKVKRQVLDPPYMKVILG 110
 QY 123 KSEPEKKEAHPIYKQLEGDGLLSLKGKNAHKKIISPFEHMKLTLVPPVLSKVTM 182
 DB 111 RSD-PKSRGSYTFVAPWIGYGLLLNQPMFOHRMLTPAFHYDILKPYGLWDSVQIM 169
 QY 183 VDKMEDKLSNGEVEVDYEWFOILTEDVISRTAF---GSSYEDGR-AVERLQAOQML-- 236
 DB 170 LDKMQLVSDSSLE--VFQDISLMTDITIKKAFSYQGSVQDLSRMSQSYIQAVGDLNN 227
 QY 237 LCAEFQKVFIPG---YFPPTRGNLKSRKLDKIRKSLKLIERRQNAIDEGSECKE 293
 DB 228 LVFAFVRNIFFQSDTIYRLSP-EGSLSHRACQLAHEDTDVIOQRKAOQEGELKVR 286
 QY 224 PAKKLLGLMTQAK-----NTVQDIYECKSPFPFAGKQTSNLTWTTLISMHPMQA 348
 DB 287 KRLDFLDVLLPARKENGSSLSDDQLRAVDVTFEBOHDTTASVSNIFALANTHEHGH 346
 QY 349 KARDEVLKVCGRDVPITDHYVKKITSLMINSRLRPVIVATIRPAKSDVKL-GGYKI 407
 DB 347 KREKEIIGLGDGASITWEHLDPMYTMCIKEMARLYPPPAISRLSDSPVFPDGRSL 406
 QY 408 PGCTLLPITLVHNDQALWGDVNERPAPFACGVPRAKHPGCFPLGAYRTICGVN 467
 DB 407 PGKFTVTLISYGLHNPNVWPP-ELFPGSRF---TOSASHNAFLPFGGARNCIGK 462
 QY 468 LAILOAKTLAVMIQRFTHLAPTYQAHAPVLMILYPOHGAPITFRSL 515
 DB 463 FANMLKVAVALTLVRFELDPFRTRIPRTARLVKLSNGIHLRLKL 510

RESULT 6
 CP3B MOUSE STANDARD; PRT; 504 AA.
 ID CP3B MOUSE
 AC 06459;
 DT 15-UTL-1999 (Rel. 38, Created)
 DT 15-UTL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cytochrome P450 3A11 (EC 1.14.14.1) (CYP11A11) (P-450UT)
 GN CYP3A11 OR CYP3A-11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ddy; TISSUE=Liver;
 RX MEDLINE=92223116; PubMed=1339292;
 RA Yanagimoto T., Itoh S., Muller-Eberhard D., Kametaki T.;
 RT "Mouse liver cytochrome P-450 (P-450IITAM1): its cDNA cloning and
 RL induction by dexamethasone." Acta 1130:329-332 (1992).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Gronow L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loguettano N.A., Peters G.J., Abramson R.D., Millaby S.J.,
 RA Bosak S.A., McEwan P.O., McKernan K.J., Malek U.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.U.M., Maria W.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC - FUNCTION: Catalyzes erythromycin 6 beta-hydroxylation.
 CC oxidation and testosterone 6 beta-hydroxylation.
 CC - CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC - oxidized flavoprotein + H(2)O.
 CC - SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC - TISSUE SPECIFICITY: Highly expressed in liver.
 CC - INDUCTION: By dexamethasone.
 CC - SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X60452; CA942981.1; .
 CC EMBL; BC010528; AAH10528.1; .
 CC PIR; S22334; A60564.
 CC HSP; P14779; JUPZ.
 CC MGI; MGI:86609; Cyp3a11.
 CC InterPro; IPR001128; Cytochrome P450.
 CC InterPro; IPR008072; EP450_CYP3A.
 CC Pfam; PF00067; P450_1.
 CC PRINTS; PR01689; EP450IICYP3A.
 CC PRINTS; PR00385; P450.
 CC PROSITE; PS00086; CYTOCHROME P450.
 CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 CC Microsome; Endoplasmic reticulum.
 CC METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 504 AA; 57854 MW; E569AF71CE23F180 CRC64;

Query Match 17.9%; Score 487.5; DB 1; Length 504;
 Best Local Similarity 28.1%; Pred. No. 4.7e-24;
 Matches 146; Conservative 93; Mismatches 178; Indels 103; Gaps 17;

QY 13 VLVLSVILSVIVKMSLMMRPKRIEHPKQSIIRGPPHFFIGNVKELVGMMLKASSH 72
 DB 4 VASLSLTVVLAISLVLYYGRKHELFKQGIIPR-----K 42
 QY 73 PMPFSNHLIPVLSFYH-NR-----KIYATPLVWGPTFRLTVADPDLREIFSK-- 123
 DB 43 PLPFL--LGYLVANVYKGLMKFDMCYKKYKGTGLPDGCPPLAIVDPETIKVNLVKEC 98
 QY 124 -SEFEKQDAPHYKQLEGDGLS-----LKGKMAHHRKISFTFEMNKLTLPVVLK 177
 DB 99 FGVFPTNRDPGV-----GIMSVAISISDDMKRYRALLSPFTSGKLEKMFVI-E 150
 QY 178 SYTDWVDMKSDLSNGEVEVDYEWFOILTEDVISRTAFSS-----YEDGRA 226
 DB 151 QYGDILVYKIRKAKGK-PTVMKDVLAGYMDVITSPGVNDSLNPPDPVEKAKK 209
 QY 227 VRLQAOQMLCAEFQKVFIPGR-----FFPFRGNLKSRLDKIRKSLKLIERRRQ 261
 DB 210 LRFDFPDLPLFSVTLFPFLTPVEMLNICMFP-----KDSIEFKKVDVYKE 258

```

QY 282 NAIDGEGECKEPPAAKDLGLMTIQAKN-----VTVQDIVECKSFPPAK 326
D 259 SRLDS-----KQKERRDPLQMLNNSHNSKDVSHKALSDWEITAGSII-----FIRAG 308
QY 327 QTSNLTMTTLLTSLWHPWOKARDEVLRVGSGSDVPTKHVYVTKLTSYLNESRLY 386
D 309 ETTSTLSFTLHSLATHPDIOCKLQDEIDALPKAPPTTDTVMEMETLDMVNLRLY 368
QY 387 PPIVATIRAKSDVYKGYKIPCGTELLIPPIAVHDDQAIWGNVNEFNPPAPDGVFRA 446
D 369 PIANRLERVCCKDVELNGVYIPKGSVTMIPSYALHDDQHW-SEPEEFQPERFSK-ENKG 426
QY 447 AKHPVGFIPFGVGTICIGONLAIIOAKLTALVMIOREPT 486
D 427 SIDPYVLLPFGNSPNCVCLQMRPLAMMKLALTKIMQNSF 466

RESULT 7
CP44_RABIT STANDARD; PRT; 506 AA.
ID P1061;
AC 01-JUL-1989 (Rel. 11, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
CY Cytochrome P450 4A4 (BC 1.14.14.1) (CYP1A4) (Prostaglandin omega-
hydroxylase) (P450-P-2).
GN CYP4A4.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NC NCBI_Taxid=9986;
RN [1]
RP SEQUENCE OF 17-506 FROM N.A., AND SEQUENCE OF 1-25.
RX MEDLINE=8007548; PubMed=3654614;
RA Matsubara S., Yamamoto S., Sogawa K., Yokotani N., Fujii-Kuriyama Y.,
RA Hanu M., Shively J.E., Gotch O., Kusunose E., Kusunose M.,
RT "CDNA cloning and inducible expression during pregnancy of the mRNA
RT for rabbit pulmonary prostaglandin omega-hydroxylase (cytochrome
RT P-450P-2)."
RL J. Biol. Chem. 262:13366-13371 (1987).
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J02818; AAA31232.1; -.
CC PIR: S32315; A29368.
CC HSSP: P14779; IUPZ.
CC InterPro: IPR001128; Cytochrome_P450.
CC Pfam: PF00067; P450.1.
CC PRINTS: PR00385; P450.
CC PROSITE: PS00086; CYTOCHROME P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
CC Microsome; Endoplasmic reticulum.
CC METAL: 453 453 IRON (HEME AXIAL LIGAND).
CC FT SEQUENCE 506 AA; 58525 MW; 50CB0412463B9B50 CRC64;

```

```

Query Match 17.9%; Score 487.5; DB 1; Length 506;
Best Local Similarity 27.2%; Pred. No. 4.7e-24;
Matches 143; Conservative 105; Mismatches 221; Indels 57; Gaps 14;

QY 16 LSVILSLVIVKSGSLMWRPKRIEER-----FSKQIGRPYHFFIGNKELVGMMLKAS 70
D 11 LSGILQVAAALGLLILLKAAQVLYHROWMLRALQCPFPFWLLGHSRE-----61
QY 71 SHMPFSSNILLPRVLSFHHKRTYGAFLVWF-GPFFRLTVADPDQIREIFSKSEFEK 129
D 62 -----FQND---QELERIQKVEKFPGACPMWSLGNKARLLVDDPDYIKILGRSD-PKA 112
QY 130 NEAHPVLVQLBDDGLSLKSEKMAHRRKLIISPTHEMLKLLVPLVKSVTDMVDKMSDK 189
D 113 PANYLMTMPWLYGYGLLLDDGQTWQHRMLTPAHYDILKRYGLAMDSVQIMLDREQL 172
QY 190 LSENGEVVDYEMFQILTEDVIGRTAF---GSSIEGRGAVFLQAOQML--LCAEAFQK 244
D 173 IQQDSLSL-IFQHSVMTLDTIMKCAFSGVQSDRNHSYIQALNDLNLVFRAR 230
QY 245 VF-----IPGYRFPPTRGNTKSRKLDKEIRKSLKLTERRONALDGESECKEPA 235
D 231 VHQSDFLYRLSPBGRLLPRAOQLAHENTDVIQ-----QRAQLQOGELEKVRKR 283
QY 296 ADLLGLMTIQAKN-----NTVQDIVECKSFPPAKQTSNLTMTTLLSMHEMOKA 350
D 284 RLDPLDVLIFAMENGSLSDQDLAEVDITMFGHDTASGVSMIFYALATHEHQRHC 343
QY 351 RDEVLRVGSRVPRKDVVVKLTLMTLSNESLRLYPPIVATIRAKSDVYL-OGYKIPC 409
D 344 RESIQGLDGDASITWEHLDQMPYTTMCIKALKLIPVPSVTQLSKPVTFPGRSLPX 403
QY 410 GTELLIPPIAVHDDQAIWGNVNEFNPPAPDGVPRAAKDPVGTIPGLGRTTIGQNTLA 469
D 404 GYILFLISLYGLHYNKWQNP-EVDPDPFRA---PDSAYSHALPFGGARGNIGKQFA 459
QY 470 IIOAKLTALVMIOREPTFLAPYQAPVLMVLPDQAPPTFRL 515
D 460 MELKVAVALTVRPELLDPFTRIPPIAVVLSKNGIHRLKRL 505

```

```

RESULT 8
CP46_RAT STANDARD; PRT; 537 AA.
ID P51871;
AC 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
CY Cytochrome P450 4F6 (BC 1.14.14.1) (CYP1F6).
GN CYP4F6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
NC NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; Tissue=Brain;
RX MEDLINE=86125358; PubMed=8554568;
RA Kawashima H., Strobel H.W.;
RT "CDNA cloning of three new forms of rat brain cytochrome P450
RT belonging to the CYP4F subfamily."
RL Biochem. Biophys. Res. Commun. 217:1137-1144 (1995).
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- TISSUE SPECIFICITY: High expression in liver and kidney. Lower
CC expression in brain.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

Query	Match	Similarity	Score	DB	Length
Best Local	Similarity 30.2%;	Pred. No. 1,1e-23;			
Matches 128;	Conservative 82;	Mismatches 182;	Indels 32;	Gaps 14	
QY	85 LSFYHWKRYGATFVWFGPTFR-LTVADDLIRIFSKSEFEYKNE--APPLVQLQEG	141			
DB	74 MGIHALGNFPDILHSWVGPPYPTLRVYHNVLAIPLOQSAVAAPKEMTLVGLKPMIG	133			
QY	142 DGLLSKGEKAAHRRKTIISPTFHMEKLLVAVVVLKSVTMDKMSDKLSNGEVEVDVY	201			
DB	134 DGLMSAGEKNNHRRLLTPAFHFPIILSKYKIKFNKSVNTHAKM-ORLTAKGSARLDWF	192			
QY	202 EMFQILTEVISRIAFG-----SSYKDGAVFLQAOQMLCAFAFOKVFIPGVRFP	254			
DB	193 EHISLWTLDSLOKCIFSPDSNQBSNRYIAILIELS---LIVKRORFVLDLPLYTL	249			
QY	255 TRGNLKRLLDKERIKSLKLLERRRQNAIDEG-EECKEPAK--DILGIMIOAKN	308			
DB	250 TADGRFRKA-CDVAVNFTDAVIFRRSRSTLNTQGVDELKRAKATKTLDFIDVLAKDE	308			
QY	309 ----VYVOIVIECKSFPAFGQQTNNLTMTTLLSHMPKQAKADEVLRYGSRDVP	364			
DB	309 HGKGLSDVDIRPADTFEMFGHDTTASLSWLTINLAHPEYQCRKQEVREILRRP-P	367			
QY	365 TK---DHVVKLTLSMILNESLRLPYATIRAKSPDKL-GSKYPCGTELLPIPIAV	420			
DB	368 EELIEMDDLAQLPLTMTICIESRLRHPVLLSRCSQDVIYDPGRVIPKGNICVISIFGV	427			
QY	421 HHQAIWGDVNEFNFPADGVPPAAKHPNGFIPFGIGVTCIGQVLAIIQAKTLAAW	480			
DB	428 HHNPISVW-PDPEVYNFRFDPENPO-KSPFLAIFIPFSGPNICIGTFAMSEIKVALALT	485			
QY	481 IORF 484				
DB	486 LIRF 489				

CC mouse liver." ;
RL Arch. Biochem. Biophys. 377:153-162(2000) .
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.
CC
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (potential).
CC
CC -1- TISSUE SPECIFICITY: Expressed in liver. Also expressed in the kidneys of female mice, with traces in the stomach, ovary, and heart of female mice and in the testis of male mice.
CC
CC -1- DEVELOPMENTAL STAGE: Detected immediately after birth in the livers of animals of both sexes, but increased with age in females; whereas it was gradually reduced in males, resulting in predominantly female-specific expression in livers.
CC
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to licenses@isb-sib.ch)
or send an email to licenses@isb-sib.ch).

DR EMBL; AB033414; BAA95951.1; -
DR MGI; MGI:1858451; Cyp3a41.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PRO1689; BP450IICYP3A.
DR PRINTS; PRO0385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KM Microsome; Endoplasmic reticulum
FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 504 AA; 57959 MW; 2B0C645BEB9CD46A CRC64;

Query Match 17.6%; Score 481.5; DB 1; Length 504;
Best Local Similarity 28.4%; Pred. No. 1,le-23;
Matches 147; Conservative 89; Mismatches 170; Indels 111; Gaps 19.

OY 18 VIISIVIVKGMGLMMRPKRKEBHFSKQIGRGPPIHYFIIGNVKELVGMLKKSSHPMPS 77
Dd :|:::|:
Db VLAIIIV---LTVRYGTTHGLFKQGIPGF-----PLPF-- 46
OY 78 HTILPVLVSFYHH-NR-----KIYGATFLWVFQPTERTLVADPDIREIFSK--SEFY 127
Dd :|:::|:
Db ---LGIVLVVYKGMLKPKMECEKKGKTWGSLFDQMPLEFVITDPENIKAVLVKCPSPVT 103
OY 47 ---LGVVLVVYKGMLKPKMECEKKGKTWGSLFDQMPLEFVITDPENIKAVLVKCPSPVT 103
Dd :|:::|:
OY 128 EKNEAHPLVKOLEGGDLIS----LKGEKAHNHRKIISPTFHNIENLKLVPVLYKSVDM 182
Dd :|:::|:
Db 104 NREEPFPV-----GIMSKAISISKDEWKRYALLSPFTSGKLKEMFPVI--EYGD I 155
OY 183 VDKRSKLSGENE--VEVDVYEWFOILTEDIYSTAGGS-----YEDGRAVER 229
Dd :|:::|:
Db 156 LVKYLMOZAKKPVTMDVLGAYSI---DVISTISGVAVDSLNPNDEPFYEKAGIDR 212
OY 230 LOAQOMILCAEAFOKVFIIPGR---FFPTRGNLKRKLDKXIRKSLKLIRRQNMI 284
Dd :|:::|:
Db 213 VDPPDFVLSVVLFPFLPVVEMLINICMF-----KOSIEFFKFVNRMESRL 261
OY 285 DGEGBECKEPAADLLGLMIOKN-----TVODIVYECKSPFPAGKQT 329
Dd :|:::|:
Db 262 DS-----KQGRVDFQLTMNHNHNSKSDSHKSLSMETTAGSIV-----FFAGYETT 311
OY 330 SNLTWTTLINSHPEWQAKADEVLAVCGSRDVPTRDYVYKLTSLMINSELALPYPI 369
Dd :|:::|:
Db 312 SSTLSFLYCLATHPIDIQCKOEIDETLPKAPFPYDTWMEYELDMVINETLRYPFG 371
OY 390 VATIRRAKSNVKGIVKIPCQTELLIPTIAVHQDAIMGNDVNEFNPARPADVPAAKH 449
Dd :|:::|:
Db 372 NLERFERKQVDELNGVYIFKSTWVISALHHDPQH--PEPEFQPERISK--ENKGSID 429
OY 450 PVGIFPGGLGVRTICQNLAIQAALKTLAVNIQRETF 486

QY 363 VPTKHVVKIKTISMLNESLRYPVATITRAKSDVKL-CGYKIPCGTELLPIIAVH 421
DB 360 SIYMHDDPPTMTGICAKLRILYPPVPGIGRELSPTFPDGRSLPGIIVLSTYGLH 419
QY 422 HDQAIWGNVDNEFNPAFADGVPRAKHPVGFIPFGLGVRTICGOMLAILQAKTLAVMI 481
DB 420 HNRKWPMP-EVTPDFRFA---PGSAQSHALPBGSGSRNCTIGQFAMNLTATLTL 475
QY 482 QRTFHLAPTYQHAPVLMILYPOGAPITTFRLTN 517
DB 476 LRFELPDPFPIPIPIARILVTKSGNGIHRLRLRFP 511

RESULT 11
CP4B RAT STANDARD; PRT; 511 AA.
ID CP4B RAT STANDARD; PRT; 511 AA.
AC P5129;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DB Cytochrome P450 4B1 (EC 1.14.14.1) (CYP4B1) (P450-isozyme 5).
GN CYP4B1 OR CYP4B-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
R3 SEQUENCE FROM N.A.
RX MEDLINE=89261667; PubMed=275471;
RA Gasser R., Philipot R.M.;
RT "Primary structures of cytochrome P-450 isozyme 5 from rabbit and rat
RT and regulation of species-dependent expression and induction in lung
RT and liver: identification of cytochrome P-450 gene subfamily IVB.";
RL Mol. Pharmacol. 35:617-625(1989).
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, M29853; AAA41778.1; -.
DR PIR; B40164; B40164.
DR HSP; P14779; IUPZ.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR OXidoreductase; MOnooxygenase; ElEctron transpOrt; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 453 453 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 511 AA; 58936 MW; CEAFBSB2ED944 CRC64;

Query March 17.6%; Score 481; DB 1; Length 511;
Best Local Similarity 27.1%; Prod. No. 1.2e-23;
Matches 144; Conservative 102; Mismatches 210; Indels 76; Gaps 20;

QY 18 VILSLVTKGMSILMWRPKIEHFSKQIGRPYHFIQNYKEL--VGNMLKASSHPMP 75

DB 20 VILMLVVLKFSLL-LRROKLAR--AMSPFGPPTNMLFGALBLQGLGSDKYVSNAAQ 76
QY 76 FSHNLLPRLVLSYHHMRKIYGATFLVMTGPTFR-LTVADPLLIEITFSKSEFFKRNANP 134
DB 77 FPH-----AHLMTFGQFVGFINIYEPDYAAVYSRGD-PRAAVYD 116
QY 135 LVKOLEGDLGLSKGEKKAHHRKIIISPTFHEMNKILVPPVILKSVTDVDCMSDKLSENG 194
DB 117 PFLQWIGKLLVLDGPKFQHRKLLTPGFHIDVLKPYAIAESTRWMLDMCKKAEENK 176
QY 195 EYEV--DYEWFPQILTEDVISRTAFG-----SSYEDGRAVFRIQAQOMLCAEAF 242
DB 177 SFDFCDVGH---VALDTMKCTFGKDGSGIGHRDNYSYLAVDLTLTMQRI---DSF 229
QY 243 Q--KVF-----PGYRFPPTRNKSLDKDKIRKSLKLERRRONAIDEGECKEPA 295
DB 230 QYHNDFTYWLTPHGRF-----LRACKIADHTDEVI---RQKALQDEKRRKIIQQ 279
QY 296 AK--DLGLMIQAKN-----VTQDIVECKSFPPAGKQTTNLTWTTLISMHPEWQA 348
DB 280 RRLDPLDILLGVDBSGIKLSDAELRAEVDTFMEGHDTTSGISWFLCMALYPEHQ 339
QY 349 KARDEYLRYCGSRDVPYTDHVVKLTLSMLNESLRYPVATITRAKSDVK-LGCVKI 407
DB 340 LCEEVVRGILGDQSPWDMLAKMTYLLTMCKECPRLYPVPQYRQLNKPFTVDSKSL 399
QY 408 PCGTELLPIIAVHDDQAIWGNVDNEFNPAFADGVPR--AAKHPGFIPLGLGVRTICG 465
DB 400 PAGSLISATLYALHNSTYW-PDEVPFLPRLS---PENAQRHPFAPMPSAGPRNCIG 455
QY 466 QNLAILQAKTLAVMIQRTFHLAPTYQHAPVLMILYPOGAPITTFRLTN 517
DB 456 QCFANMEKRVYALCLHREFSLDPSKMPKIVQILRSKNGIHLLKPLAS 507

RESULT 12
YRV8 CAEEL STANDARD; PRT; 519 AA.
ID YRV8 CAEEL STANDARD; PRT; 519 AA.
AC Q27520;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DB Putative cytochrome P450 CYP13A1 (BC 1.14.-.-).
GN CYP13A1 OR T1089.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
R3 SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RA Gardner A.;
RT Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. They oxidize a variety of structurally unrelated
CC compounds, including steroids, fatty acids, and xenobiotics.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z48717; CA488410.1; -.
DR PIR; T24784; T24784.
DR HSP; P14779; IUPZ.
DR WormPep; T1089.8; CE01660.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.

DR PRINTS: PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Hypothetical protein; Oxidoreductase; Heme.
 FT METAL 465 465 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 519 AA; 60044 MW; 3AF37EDC43539D7A CRC64;
 Query Match 17.6%; Score 480; DB 1; Length 519;
 Best local similarity 28.5%; Pred. No. 1.5e-23;
 Matches 144; Conservative 86; Mismatches 182; Indels 94; Gaps 19;
 QY 32 WWRPKLIEHPSKQGIKP--PYHFFIGNVKELVGMMLKASHMPESEHNLPRVLSFY 88
 DB 26 FWR-----RGGVVGPMGFP-----VLGVFLNSLDNNEFPFLQOC----- 58
 QY 89 HHMKITGATFLVWFGFTPLTVADPDILREIF--SKSEFEKNEAHPLVKQLEGD--- 142
 DB 59 REMWRKFGKLYGFTGGLKTLVSDPLVHEVVOYDINFGKR-----NFIQDSKE 113
 QY 143 ---GLSLKGEKMAHRRKISPTFHENMLKLVPVYLKSVTMDKSDKSENSEVEVD 199
 DB 114 KRTLPFAAGFRWKLALISSPTNSGLAKLYQVEDSALELL--RHIEKQAGGK-QID 171
 QY 200 VYEWPIITEDVIRITLFGSSYEDGRAVFLQAOQML--LCAEAFQKVIIPGKFFPT-- 255
 DB 172 MLKPYQEFITLDVIRIMG--QDSQMFKNPIPIVSKLPQGNFALFLIG-GIFPTFL 227
 QY 256 -----RGMLK--SRKLDKEIRKSLKLIERRQNAIDGEGECKEPPAKDILGLMI 304
 DB 228 VEITRQILLKMLKGSFRKINETLIDAHIRIKOREED--QKNGIIGRPA--DFIDFL 283
 QY 305 QAK-----NVTVDIVECKSPFPAGKQTTSNLTLTWTTILS 341
 DB 284 DAKEDVHEFGENNDGFSKTYTNRQLTTEIEIVGQCTVFLIGPPTLISLSTAYTLA 343
 QY 342 MHPWQAKARDEVLRVCGSRDVPKDHVVLKTLMSLINESLTP--PIVATIRAKSDV 400
 DB 344 THPEIQKLQEEVNRCEPNPEV--TIDQSLKLTKECVFKALMLYLGLAFNSRRCMT 402
 QY 401 KLGVKIPCGTELLPIIAVHDDQAINGVNENFPAFADGVPPAKKPVGIFPGGLCV 460
 DB 403 KLGNMKVVGTMIDVDTWTLHTDPNIGDADDEPKPERWQTPNSDQIYKSGVPIPGAP 462
 QY 461 RTICQNTALIQAKTLTAVMIGRTFF 486
 DB 463 RQIGIMRLAYMEKILLVHLIRKPTF 488
 RESULT 13
 CP3G_MOUSE STANDARD; PRT; 504 AA.
 ID CP3G_MOUSE
 AC Q64481;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 3A16 (EC 1.14.14.1) (CYP3A16).
 GN CYP3A16 OR CYP3A-16
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR; TISSUE=Liver;
 RX MEDLINE=95112853; PubMed=7813478;
 RA Itoh S.; Seton W.; Abe T.; Hashimoto H.; Yanagimoto T.; Kametaki T.;
 RT "A novel form of mouse cytochrome P450 3A (Cyp3a-16). Its cDNA cloning
 and expression in fetal liver.";
 RL Eur. J. Biochem. 226:877-882(1994).
 CC FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. In liver microsomes, this enzyme is involved in an
 CC NADPH-dependent electron transport pathway. It oxidizes a variety
 CC of structurally unrelated compounds, including steroids, fatty
 CC acids, and xenobiotics.

CC -| CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -| SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -| DEVELOPMENTAL STAGE: Fetal- and puberty-specific.
 CC -| INDUCTION: P450 can be induced to high levels in liver and other
 CC tissues by various foreign compounds, including drugs, pesticides,
 CC and carcinogens.
 CC -| SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or-send-an-email-to-licence@isb-sib.ch)
 CC or send an email to licence@isb-sib.ch).
 CC
 CC EMBL; D26137; BA05133.1; -
 CC PIR; S50892; S50892.
 CC HSBP; P14779; IUPZ.
 CC MGD; MGI:106099; Cyp3a16.
 CC InterPro; IPR001128; Cytochrome P450.
 CC InterPro; IPR008072; BP450_CYP3A.
 CC Pfam; PF00067; P450; 1.
 CC Pfam; PF01689; BP450IICYP3A.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 504 AA; 57929 MW; E364B91FD10D5D1A CRC64;
 Query Match 17.6%; Score 479.5; DB 1; Length 504;
 Best local similarity 28.2%; Pred. No. 1.5e-23;
 Matches 145; Conservative 86; Mismatches 175; Indels 107; Gaps 18;
 QY 18 VILSLVVKQMSILMWRPKIEHPSKQGIKRPYHFFIGNVKELVGMMLKASHMPEPS 77
 DB 13 VLAILIV-----LLYVGYTTHGLFKKQIGIPG-----KPLPF- 46
 QY 78 HNLPRVLSYTH--WR-----KYGATFLVWFGFTPLTVADPDILREIFSK---SEFY 127
 DB 47 ---LGLVANYKGLMFWDECEYKGTGLFPGQIPLFVITDPEITIKAVLYKCSRVT 103
 QY 128 EKQEAHPLVYKQSDGLS-----LGEKMAHRRKISPTFHENMLKLVPVYLKSVTDM 182
 DB 104 NRQDFEV-----GIMSKSISLADDEMKRYRLSLPTFSGNLKMFPVI--EQYDI 155
 QY 183 VDKWSDLSNGEVEVDVYEWPIITEDVIRITLFGSSYEDGRAVFLQAOQML--LCAEAFQKVIIPGKFFPT-- 255
 DB 156 LVKYLROBAEKGK--PVAVKDVLGAYSMVDVITTFGVNDISLNNPEDPVENAKRYLAPD 214
 QY 222 AQQMLLGAEPQKVFPGVR-----FPTFGNLSKSKLDEIRKSLKLIERRQNAIDG 286
 DB 215 YFDPLSLVALFPPLPIYMLNICMF-----KDSIEFPKTVDMTEKRLDS 263
 QY 287 EGEECKEPPAKDILGLMIQAKN-----VTVDIVECKSPFPAGKQTTSNLTLTWTTILS 341
 DB 264 -----KQKHRDPIYLWMEAYNKSQKDSHKLSEIEITAQSIIT-----FIFAGYPTSS 313
 QY 332 LLTWTTILSNHPWQAKARDEVLRVCGSRDVPKDHVVLKTLMSLINESLTP--PIVIA 391
 DB 314 ILSTVYSLATHPIQKLQEEIDEALPNPAPPTTYTAVAMEITDMLVNETLALYITNR 373
 QY 392 TIRPAKSPVKGKYGKIPCGTELLPIIAVHDDQAINGVNENFPAFADGVPPAKKPV 451
 DB 374 LQVCKQKQVEINGIYIKGSGTVIIPSVLHDDQHM--PEDEEOPRFSK--ENKGSIDPY 431
 QY 452 GFIPPGGLGVRCIQNTALIQAKTLTAVMIGRTFF 486
 DB 432 YLPPGNGPRNCIMRFLMMKRLALIKVLQNSF 466

RESULT 14
 YRV2_CABEL STANDARD; PRT; 520 AA.
 ID YRV2_CABEL
 AC Q27514;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative cytochrome P450 CYP13A5 (EC 1.14.-.-).
 GN CYP13A5 OR T10B9.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCB1_TaxID=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=bristol N2;
 RA Gardner A.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. They oxidize a variety of structurally unrelated
 CC compounds, including steroids, fatty acids, and xenobiotics.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; 248717; CAA88604.1;
 DR PIR; T24778; T24778.
 DR HSSP; P14779; JUPZ.
 DR WormPep; T10B9.2; CE01656.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 DR HYPOTHETICAL protein; Oxidoreductase; Monooxygenase; Heme.
 FT METAL 464 464 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 520 AA; 59524 MW; 0B7B19E2587ADF3B CRC64;
 Query Match 17.5%; Score 478; DB 1; Length 520;
 Best Local Similarity 25.4%; Freq. No. 28-23; Indels 128; Gaps 21;
 Matches 143; Conservative 109; Mismatches 183;
 QY 23 VIYKMSL-----LWRRPKIEEHFKQIRGPPYHFGVYKELVGMKAKASHMP 75
 DB 5 ILTAGSFTGLTYIYW-----IMSWIRKGVKGRGPFPEVHE----- 45
 QY 76 FSHNLPVLSFYTHRRKIYGATFLVWFPFTRLVADPDILREIFSK--SEFEKNEAH 133
 DB 159 LDDASAKGKA-VDLIDYQEFITDIIGRIAMGQTESLMFRNPMLEPKVGIFDGRKLPFL 217
 QY 231 QAQOMLILCAEAPKQVI-----PGRAFPTRGNLKSRKLDREIKRSILKLIERR--RQ 281
 DB 218 VSGIFPLAGTWEEFFRRFPFSIOPADIMST-----VAKALKKLEQAADEK 265
 QY 134 PLVKOLEGD-----GLSLKGEKVAHRKLIISPTFHEMLKLVPAVLSVTDVYKM 186
 DB 102 --TNPIQADBNKKRAHLVSAQHRKRLRTSSPFSKRLKRLKMSVVEEYVELM-RH 158
 QY 187 SDKLSENGEYEVVDVYEMFOILTEDVISRTAFSS-----YEDGRAVRL 230
 DB 159 LDDASAKGKA-VDLIDYQEFITDIIGRIAMGQTESLMFRNPMLEPKVGIFDGRKLPFL 217
 QY 231 QAQOMLILCAEAPKQVI-----PGRAFPTRGNLKSRKLDREIKRSILKLIERR--RQ 281
 DB 218 VSGIFPLAGTWEEFFRRFPFSIOPADIMST-----VAKALKKLEQAADEK 265
 QY 282 NAIDGEGECKEPAANDLGLMTQA-----KNTVODIYEE 317
 DB 266 AGEIPSGE-----PODFIDLFIDARANDVFEESALGFAKTEIAYDKOLTFDEIIGQ 319
 QY 318 CKSFFAGKQTTNLTWITLILSMDEMAKRDVNLVCGSRDVPFTDHYVKKLTLSM 377

DB 320 LVEFLLAGYDTALSLSYSSYLLARHPEIOKLUOEYDRECNPEV-TFDQISKLYMEC 378
 QY 378 ILNESLRYPPIVATI--BRAKSDVILGGYKIPCCGTELLIPILAVHDDAIVGNVNEF 434
 DB 379 VKKALRYNP--LASTVHNKRMKEETNVLDVGQLEKGNVGVDDTWLTALYDPRVGEDANEF 436
 QY 435 NPAPADGVPPAAHVPVGFIPFGIVTTCIGONLAILAQCLTAVMQRTEH----- 487
 DB 437 RPERWESG-DELFYAKGGVLPFGWGPRICTGMRLAWMEKMLTLIKRYTFSTQTEI 495
 QY 488 ---LAPTYOHAP-TVLMILYPOH 506
 DB 496 PLKLVGSAKTTAPRSVMLKLTIPRH 518
 RESULT 15
 CP4B_HUMAN STANDARD; PRT; 511 AA.
 ID CP4B_HUMAN
 AC P13584;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 4B1 (EC 1.14.14.1) (CYP1B1) (P450-HB).
 GN CYP4B1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCB1_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=90105307; PubMed=2574990;
 RA Nhaburp P.T.; Gonzalez F.J.; McBride O.W.; Gelboin H.V.; Kimura S.;
 RT "Identification of a new P450 expressed in human lung: complete cDNA
 RL sequence, cDNA-directed expression, and chromosome mapping.";
 RN Biochemistry 28:8060-8066(1989).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=90126824; PubMed=2298205;
 RA Yokotani N.; Sogawa K.; Matsubara S.; Gotch O.; Kusunose E.;
 RT "cDNA cloning of cytochrome P-450 related to P-450p-2 from the cDNA
 RL library of human placenta. Gene structure and expression.";
 RL Eur. J. Biochem. 187:23-29(1990).
 CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. In liver microsomes, this enzyme is involved in an
 CC NADPH-dependent electron transport pathway. It oxidizes a variety
 CC of structurally unrelated compounds, including steroids, fatty
 CC acids, and xenobiotics.
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -1- INDUCTION: P450 can be induced to high levels in liver and other
 CC tissues by various foreign compounds, including drugs, pesticides,
 CC and carcinogens.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; J02871; AAA5712.1; -
 DR EMBL; X16699; CAA34672.1; -
 DR PIR; S07765; C4HDB1.
 DR Genew; HSNCL2644; CYP4B1.
 DR MIM; 124075; -
 DR GO; GO:0015034; F:cytochrome P450 activity; TAS.
 DR InterPro; IPR001128; Cytochrome_P450.

DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT METAL 453 453 IRON (HEME AXIAL LIGAND).
 FT CONFLICT 37 37 R -> Q (IN REF. 2).
 SQ SEQUENCE 511 AA; 59019 MW; E24410DF0707A8F6 CRC64;

Query Match 17.5%; Score 477; DB 1; Length 511;
 Best Local Similarity 27.8%; Pred. No. 2,2e-23;
 Matches 148; Conservative 92; Mismatches 190; Indels 102; Gaps 22;

```

QY 10 IPKVLVLS-----VILSLIVKMSLWMPRKIEHFSKQIGRPYHFFIGNV 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 VPSFLSLSPSSLSIGMASGLIVGLFKLHL-LRRRTLAKAMDK--PPGPTHWLFQHA 58
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 60 KEL--VGMMLKAS--HPMPSHNLPRVLSFYHHKRIYGAFLVMGPTFR-LTVADP 114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 59 LEIQTGSLDKVSMHQFPYAHPL-----WFGQFTGFLNTIYEP 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 115 DLIREFSKSEFYKNAHPLVVKQLEGDLISLGEKMAHRRKISPTFMENLKLIVPV 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 98 DYAAVYSRGD-PKAPVYDFFLWIGRGLVLEGPKMLQHRKLLPGFHYDLKPYAV 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 VLKSVTDVMDKMSDKLSENGEVEV--DYEMFQILTEDIYSTAFG-----SSTE 222
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 157 FTETRIALDKWEKAREGKSPDIFCDVGH---MALNTLMKCTFGRGDTGLGHRDSSY 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 223 DGRAVFLQAQOMLCAEAFQKVEI-----PGYRF-----PTRGNLKSRL--DK 266
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 213 LAVSDLTLMQQLVSTQ-YHNDFTYLTGGRFLPACQVAHDHTDQVIRERKALQDE 271
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 267 EIRKSLKLTERRONALDGESECKEPAKDLLGLMIQAKN---VTVDIVBECKSF 321
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 272 KVRKK---IQNRH-----LDFILDLGARDEDIDKLSADLRAEVDTF 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 322 FPAKQTSNLLTTTLLSMHPWQAKARDEVLRVCGSRDVPYTKDHYVKLTLSMILNE 381
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 313 MEEGHTTSGISWFLYCMALYPHQHRCREVEIREIDGFQWDDLGKMTYLLTMCIE 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 382 SLRLVPPIVATIRRAKSDVK-LGGYKIPCGTELLPIIAVHDAQMGNDVNEFNAPPA 440
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 373 SFRILPVPQYRQLSKPVTVFDGRSLPAGSLISMHIYALHRNSAVW-PDPEVFDSLRF 431
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 441 DGVPRAAK-HPVGFIPEGLGVRTICIGNLAILQAKLTAVMIQRTFTHLAPT 491
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 432 --TENASRHHFAFPFAGPRNCIGQAFMSEMKVVTAMCLLRFPSLDPS 481
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: May 19, 2004, 10:26:20
 Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 10:20:55 ; Search time 46 Seconds
(without alignments)

3566.728 Million cell updates/sec

Title: US-09-992-901-2

Sequence: 1 MBESSSWFKPKVLSVYL.....LTYFGGAPITFRRLTNHED 520

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_ricent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2722	99.7	520	10 Q48786	Q48786 arabidopsis
2	1673	61.3	589	10 Q84SB7	Q84SB7 oryza sativ
3	1404	51.4	537	10 Q8LIF2	Q8LIF2 oryza sativ
4	1187	43.5	528	10 Q8LJ74	Q8LJ74 zea mays su
5	1167.5	42.8	527	10 Q8LGM8	Q8LGM8 zea mays su
6	1165.5	42.7	523	10 Q9PDD1	Q9PDD1 oryza sativ
7	1154.5	42.3	525	10 Q9ATU3	Q9ATU3 lolium rigi
8	1151.5	42.2	525	10 Q9ATU2	Q9ATU2 lolium rigi
9	1147.5	42.0	525	10 Q9ATU4	Q9ATU4 lolium rigi
10	1145.5	41.7	525	10 Q9ATU5	Q9ATU5 lolium rigi
11	1138.5	41.2	519	10 Q9AX23	Q9AX23 oryza sativ
12	1126	41.2	515	10 Q9JUD2	Q9JUD2 arabidopsis
13	1124.5	41.2	515	10 Q8M1E1	Q8M1E1 arabidopsis
14	1124.5	41.0	528	10 Q8LGN4	Q8LGN4 zea mays su
15	1119	41.0	524	10 Q42700	Q42700 catharanthu
16	1113	40.8	524	10 Q42700	Q42700 catharanthu

17	1112	40.7	511	10 Q9FED1	Q9FED1 oryza sativ
18	1111	40.7	531	10 Q9FED3	Q9FED3 oryza sativ
19	1110	40.7	516	10 Q42701	Q42701 catharanthu
20	1110	40.7	528	10 Q8S9F0	Q8S9F0 oryza sativ
21	1108	40.6	528	10 Q9PE60	Q9PE60 oryza sativ
22	1098	40.2	512	10 Q9LUD3	Q9LUD3 arabidopsis
23	1089.5	39.9	512	10 Q9LUC6	Q9LUC6 oryza sativ
24	1088.5	39.9	531	10 Q9PEP3	Q9PEP3 oryza sativ
25	1080.5	39.6	535	10 Q8LIC7	Q8LIC7 arabidopsis
26	1079.5	39.5	512	10 Q9LUC8	Q9LUC8 oryza sativ
27	1079	39.5	544	10 Q9FEB1	Q9FEB1 arabidopsis
28	1075.5	39.4	512	10 Q9LUC5	Q9LUC5 oryza sativ
29	1067	39.1	565	10 Q9PEE6	Q9PEE6 oryza sativ
30	1060	38.8	552	10 Q8LIC6	Q8LIC6 oryza sativ
31	1050.5	38.5	505	10 Q9LUD0	Q9LUD0 oryza sativ
32	1047	38.4	532	10 Q9PE66	Q9PE66 oryza sativ
33	1045	38.3	498	10 Q9M4X2	Q9M4X2 lycopersico
34	1040.5	38.1	519	10 Q8LHV0	Q8LHV0 oryza sativ
35	1028.5	37.7	505	10 Q9PEK4	Q9PEK4 arabidopsis
36	1024.5	37.5	512	10 Q9LUC9	Q9LUC9 arabidopsis
37	1023	37.5	508	10 Q40411	Q40411 nicotiana p
38	1021.5	37.4	520	10 Q8L4Q4	Q8L4Q4 oryza sativ
39	1019	37.3	517	10 Q80729	Q80729 arabidopsis
40	1017.5	37.3	513	10 Q8S9C0	Q8S9C0 solanum tub
41	1015	37.2	517	10 Q8LIR5	Q8LIR5 oryza sativ
42	1006	36.8	534	10 Q94EV6	Q94EV6 zea mays (m
43	999	36.6	519	10 Q9ASR3	Q9ASR3 arabidopsis
44	997	36.5	430	10 Q8LJ72	Q8LJ72 arabidopsis
45	990.5	36.3	518	10 Q9FRC3	Q9FRC3 oryza sativ

ALIGNMENTS

RESULT 1
ID Q48786 PRELIMINARY; PRT; 520 AA.
AC Q48786;
DT 01-UN-1998 (TREMBLrel. 06, Created)
DT 01-UN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-UN-2003 (TREMBLrel. 24, Last annotation update)
DE Putative cytochrome P450.
GN AT2G26710.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.B., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen W., Vanaken S.B., Mayam L., Talon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana".
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SMITARIITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AC003105; AB95305.1; -.
DR FIC; H84663; H84663.
DR HSP; P14779; IBVY.
DR GO; GO:000497; P:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 DR Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 520 AA; 59490 MW; 585186A0CC07A59 CRC64;

Query Match 99.7%; Score 2722; DB 10; Length 520;
 Best Local Similarity 99.8%; Pred. No. 1.8e-216;
 Matches 519; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESESSSWPIPKLVLSVLSVLYVKGMSLWMPRKIEEHSKGGIRGPPYHFIQV 60
 DB 1 MESESSSWPIPKLVLSVLSVLYVKGMSLWMPRKIEEHSKGGIRGPPYHFIQV 60
 QY 61 ELVGMMLKASHMPFSNHLPRVLSFYHMKRIYGAFTLVWFGPTFLVADPDIREI 120
 DB 61 ELVGMMLKASHMPFSNHLPRVLSFYHMKRIYGAFTLVWFGPTFLVADPDIREI 120
 QY 121 FSKSEFEKNEAHPVLYKQLEGDGLSLKGEKNAHRRKISPTFHENIKLVVPLKSVT 180
 DB 121 FSKSEFEKNEAHPVLYKQLEGDGLSLKGEKNAHRRKISPTFHENIKLVVPLKSVT 180
 QY 181 DWVDKMSDKLSENGEVDVYEWFOITLTDVTSRTAFSGSYEDGAVFRLQAQMLCAE 240
 DB 181 DWVDKMSDKLSENGEVDVYEWFOITLTDVTSRTAFSGSYEDGAVFRLQAQMLCAE 240
 QY 241 AFQKVPFGYRFPPTNGNLKSKRLDYKIKSLIKLIERRQNALIGSECEKPAKDIL 300
 DB 241 AFQKVPFGYRFPPTNGNLKSKRLDYKIKSLIKLIERRQNALIGSECEKPAKDIL 300
 QY 301 GLMIOAKNTVODIVECKSFPAFGKQTSNLTWTTLISMPHMOAKADEVLRVCGS 360
 DB 301 GLMIOAKNTVODIVECKSFPAFGKQTSNLTWTTLISMPHMOAKADEVLRVCGS 360
 QY 361 RDVPTKHVVVKKLTSLMILNESLRILYPIVATIRAKSDVKGKIPGTELLIPITAV 420
 DB 361 RDVPTKHVVVKKLTSLMILNESLRILYPIVATIRAKSDVKGKIPGTELLIPITAV 420
 QY 421 HHQOATWGNVNEFNPAFADGVPRAAKPVGFIPIGIGVKTICIGNALILQAKTLAVM 480
 DB 421 HHQOATWGNVNEFNPAFADGVPRAAKPVGFIPIGIGVKTICIGNALILQAKTLAVM 480
 QY 481 IORTFHIAPTYOHAPTVMLLYPOGAPITFRRLTNHD 520
 DB 481 IORTFHIAPTYOHAPTVMLLYPOGAPITFRRLTNHD 520

RESULT 2
 Q84SB7 PRELIMINARY; PRT; 589 AA.
 AC Q84SB7;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE OSJNB0008D07.14 protein.
 GN OSJNB0008D07.14
 OS Oryza sativa (japonica cultivar-group).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Eubacteriobacteriota; Oryzaeae; Oryza.
 CC NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RC MEDLINE=22337376; PubMed=12447438;
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
 RA Hosokawa S., Maekawa M., Arikawa K., Chiden Y., Hayashi M.,
 RA Okamoto S., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
 RA Hijiwata S., Honda W., Ichikawa Y., Ikonuma A., Iijima M., Ikeda M.,
 RA Ikono M., Itoh S., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,

RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
 RA Shimokawa T., Shomura A., Song J., Takasaki Y., Terasawa K., Tsuji K.,
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
 RA Yano M., Jiang J., Gojobori T.,
 RT "The genome sequence and structure of rice chromosome 1."
 RL Nature 420:312-316(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Katayose Y.,
 RA Submitted (Feb-2003) to the EMBL/Genbank/DBJ databases.
 DR EMBL:AB006237; BAC57879.1;
 DR GO:0006118; P:electron transport; IIA.
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 SQ SEQUENCE 589 AA; 66070 MW; 23892D1FB457DF4B CRC64;

Query Match 61.3%; Score 1673; DB 10; Length 589;
 Best Local Similarity 56.6%; Pred. No. 1.3e-129;
 Matches 318; Conservative 87; Mismatches 79; Indels 78; Gaps 8;

QY 23 VIVKMSLWMPRKIEEHSKGGIRGPPYHFIQVLYVGMMLKASHMP--FSHN 79
 DB 17 VAVKLVVWMPRKIEEHSKGGIRGPPYHFIQVLYVGMMLKASHMP--FSHN 76
 QY 80 ILPRVLSFYHMKRIY----- 95
 DB 77 VLPRVLSFYHMKRIYGNPPPPPLNLNLSLQKQOPRTNRQVAVGERRAPGSDIDM 136
 QY 96 ---GATFVWPGPTFLVADPDIREI-FSKSEFEKNEAHPVLYKQLEGDGLSLKGEK 151
 DB 137 AALGFTFLWPGPTFLVADPDIREI-VLPRVLSFYHMKRIYGNPPPPPLNLNLSLQKQOPRTNRQVAVGERRAPGSDIDM 136
 QY 152 MAHRRKISPTFHENIKLVVPLKSVTDVWDK---SDKSENGEVDVYEWFOI 206
 DB 197 MAHRRVLPFAFHMNLRLLPVGMFTVDMADKRRANABDK--SEVEIDVSDWFOV 253
 QY 207 IREDDVSRAPSSVEDGAVFRLQAQMLCAEAFQVFIPIGFRFPPTNGNLKSKRLDK 266
 DB 254 VTEIDALTRAFGRSIEDGVVFKLQOLMARASAFKRVFIPIGFRFPPTNGNLKSKRLDK 313
 QY 267 EIKSLIKLIERRQNALIGSECEKPAKDILGMLTQA-----KNTVQ 312
 DB 314 EIKNLVTLIGRQEGDDEKLDG---AKDLGLMILNAAASSNGRSALPVSPITVN 369
 QY 313 DIVECKSFPAFGKQTSNLTWTTLISMPHMOAKADEVLRVCGSDVPTKHVVYL 372
 DB 370 DIVECKTFPAFGKQTSNLTWTALVLMHFEWERKQEVLDVCGADGVPSRQLAKL 429
 QY 373 KTLMSILNESLRILYPIVATIRAKSDVKGKIPGTELLIPITAVHDDQATWGNV 431
 DB 430 KTLGMLINTIRILYPIVATIRAKSDVKGKIPGTELLIPITAVHDDQATWGNV 489
 QY 432 NENNPAPFADGVPRAAKPVGFIPIGIGVKTICIGNALILQAKTLAVMIRFTHAPT 491
 DB 490 AQPNPAPFADGVPRAAKPVGFIPIGIGVKTICIGNALILQAKTLAVMIRFTHAPT 549
 QY 492 YOHAPTVMLLYPOGAPITFR 513
 DB 550 YVHAPTVMLLHPQYCAPIVFR 571

RESULT 3
 Q84IF2 PRELIMINARY; PRT; 537 AA.
 ID Q84IF2;
 AC Q84IF2;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, last annotation update)
 DE Putative cytochrome P450.
 GN OJ1316_A04.10 OR P0503D09.27.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrharioideae; Oryzaceae; Oryza.
 NCBI_Taxid=39947;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 clone: OJ1316_A04.10";
 RN Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
 clone: P0503D09.27";
 RN Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AP003822; BAC06993.1; -.
 DR EMBL; AP005455; BAC16737.1; -.
 DR Gramene; OBLIR2; -.
 DR GO; GO:0004553; F:hydroxylase activity, hydrolyzing O-glycosyl . . .; IEA.
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome P450.
 DR InterPro; IPR008263; Glyco_hyd16_AS.
 DR Pfam; PF00067; P450.1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450.1.
 DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
 DR Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 537 AA; 59909 MW; 068441D7D046ED20 CRC64;

Query Match 51.4%; Score 1404; DB 10; Length 537;
 Best Local Similarity 51.1%; Pred. No. 2e-107;
 Matches 271; Conservative 97; Mismatches 138; Indels 24; Gaps 10;
 QY 7 SNFIPKVLVLSVLSLVV-----KMSLWMPRKIEHFSKQIGRPYHFFIGNVKE 61
 DB 10 SMSGAAVAVAAMAAVVAAMAAALWMPRRVRHPPAAGVGPGVRFVSSIE 69
 QY 62 LVGMMLKASSHM--PESHNLPRVLSFYHMKIKYGATLVWGFRTLVADPDLIRE 119
 DB 70 LVRLVNDASREMEPTSHDILPRVLPFYHMRKLYGKMLIWGRTPRLVSPDLIRE 129
 QY 120 I-FSKSEFEKNEAPLVKOLEGDLISLKGKNAHRIKISPTFHENIKLVYVVLKS 178
 DB 130 VLLTRADHDHREARPMICPEGYLSLHGERNARRRRLTPAFHENEIMLAFVAGT 189
 QY 179 VTDVYDKMSDKSEN--GEVEVDYEMQIITEDVISRTANG--SSYEDGAVRLDAQM 225
 DB 190 VTRMDDELAEARAGAGAEVDVEMFORVDEAITFAAGRRNYDDGAVRLDQELA 249
 QY 226 LICAFAQVFIPIGRFFPTRGNLKSRKLDKEIRKSLKLIERRQNALIGESECKE-- 293
 DB 250 GYATNAHSKVYIPGRFLPRTKRRVWQLDREIRSLAKFV--TGLQSSSHDDADDG 308
 QY 294 ---PAKDLGLMIOAKVTVVDIVECKSPFPAKQTTSNLITWTTTILISHPENQAKA 350
 DB 309 DGGGGRBEMSMFAA--MTAGEIIEESKNEFFAKETLSNLTWTTVALAMPWQERA 366
 QY 351 RDEVLRVCGSRDVPRTKDVVVKLTSMILNESLRLPYIVATIRASDVYLGQYKPCG 410
 DB 367 RKEVAVVCGRQDLPTRKDLPKLKTIGMLINETLRYPYAVAMITAEDEVLEGGCVAPAG 426
 QY 411 TELLPITAVHHDAQIWNQVNEFNPAFA--DGVPAAGVPGVIFPGLGVYTCIGON 467

DB 427 TEWPIPIAVVHHDAAMGDDAAEFNPARFAADDGGR--RHPMAFMREGGARVCIGON 484
 QY 468 LAIOAKITLAWMIQRTFFHLPAYQAFLVLMILYPOHGAFTTFRLLTN 517
 DB 485 MALMEAKVALVLRERFERLSPAYVAPRVLMLTSPQGAIVIFRPLTS 534
 RESULT 4
 ID OBL174 PRELIMINARY; PRT; 528 AA.
 AC OBL174;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Cytochrome P450 monooxygenase CYP72A16.
 OS Zea mays subsp. mays (maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoidae; Andropogoneae; Zea.
 NCBI_Taxid=4578;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Indred B73;
 RA Wang U., Schuler M.A.;
 RT "Molecular Characterization of the Maize CYP1C3 and CYP72A subfamily
 genes";
 RN Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Indred B73;
 RA Wang U., Schuler M.A.;
 RT "Xenobiotics Regulate Maize Cytochrome P450 Monooxygenases at a Post-
 Transcriptional Level";
 RN Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF455265; AA077716.1; -.
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome P450.
 DR Pfam; PF00067; P450.1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450.1.
 DR Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 528 AA; 60642 MW; 95046F1C7112CCD CRC64;

Query Match 43.5%; Score 1187; DB 10; Length 528;
 Best Local Similarity 43.8%; Pred. No. 1.8e-89;
 Matches 229; Conservative 99; Mismatches 177; Indels 18; Gaps 4;
 QY 4 ESSSFPIKVLVLSVLSLVVKMSLWMPRKIEHFSKQIGRPYHFFIGNVKELV 63
 DB 13 EASPARAGATAAVLVVLAAMTLENAMWTPRLDALPAQGLKGRYLLTGDVRENA 72
 QY 64 GMLLKASSHMPF--SHNILPRVLSFYHMKIKYGATLVWGFRTLVADPDLIREIFS 122
 DB 73 RLNREARKPLPLSHDIIPRVLPFNAYKENGNTSFTWFGRLPYIIPDLPELMEVLS 132
 QY 123 KSEFEKNEAPLVKOLEGDLISLKGKNAHRIKISPTFHENIKLVYVVLKSVTM 182
 DB 133 NKFGFGKPLPSRSGKLANLANHBEKAKNHRILNPFHNEKIGMLVYATGCCADM 192
 QY 183 VDKMSDKLSENGEVEVDVYEFQILTEDVISRTAFSSSYEDGAVRLDAQMILCAEAF 242
 DB 193 INRWENSSSKPEBMDVWPFQVLTGDVISRTAFSGNYDEGRNIFQLQEQAEIRLIQSF 252
 QY 243 QKVPFPGYRFPPTGNLKSRLDKEIRKSLKLIERRQNALIGESECKEPAKDLITG 302
 DB 253 QTFIPGYWFLPTKNNRMEKIDEIRKILHGIIRKKEAFIDSEGN-----DDLGS 306
 QY 303 MIOAK-----NTVVDIVECKSPFPAKQTTSNLITWTTTILSNHPENQAKARD 352
 DB 307 LVESNMRSSNNAKLGWTTEDILIECKLTFYAGGETTSVLLTWTLLLSNHPENQCAR 366

QY 353 EYLKVGSGRDVPTKDHVVKLTSLMINESLRLYPPIVATIRRAKSDVKLGKYPICGTE 412
 DB 367 EVLNHNG-MGTPDPDLNRLKVTMLLYVLAFLYPPVFLSKRTYKEMELGKIKPSCVS 425
 QY 413 LLIPITIAVHHDAIWNADVNEFNPAFADGVPRAAKHVGFIPFGLSVRTICIGONLALIQ 472
 DB 426 LLIPITIFIHDPINWCKDASEFNPQRFEDGINSATGHQAFFPFGMGPRICIGONFALLE 485
 QY 473 AKLTAVMIOREFFHLAPTYOAHPTVLMLYPOHGAFTFRRL 515
 DB 486 AKALSTILQRFSELSSTYTAHYVTTLHPHQAQIRLKL 528

RESULT 5

Q8LGM8 PRELIMINARY; PRT; 527 AA.
 AC Q8LGM8; 01-OCT-2002 (TEMBLrel. 22, Created)
 DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE Cytochrome P450 monooxygenase CYP72A5.
 OS Zea mays subsp. mays (maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 ON NCBI_TaxId=4578;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Inbred B73;
 RA Wang J., Reeds R., Schuler M.A.;
 RT "Molecular characterization of the maize CYP71C3 and CYP72A subfamily
 genes.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Inbred B73;
 RA Wang J., Reeds R., Schuler M.A.;
 RT "Xenobiotics regulate maize cytochrome P450 monooxygenases at a post-
 transcriptional level.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AY072300; AAL66770.1; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 527 AA; 60461 MW; 2271A93503F1F912 CRC64;

Query Match 42.8%; Score 1167.5; DB 10; Length 527;
 Best Local Similarity 42.8%; Pred. No. 7,2e-88;
 Matches 226; Conservative 102; Mismatches 180; Indels 19; Gaps 5;

QY 1 MEEBSGMPFKVYLVSILSVIYKGMSLMWRPKRIEHSKQGINRPPHYFIGNVK 60
 DB 8 MLEKVSPPMALASV-VASVSLMLVMTLEWAMTPTLRRLARVQLKGTFRRLPTGLDR 66
 QY 61 ELVGMALKASGHPMPF-SNNILRVLSFYHNRKIYGATFLVFGTFPLVADDLIRE 119
 DB 67 ETARANREARKKPLPLGSHDIAPRVPMHSHSTKEGKSLFTWFGTFPMVLPDELVE 126
 QY 120 IFSKSEFYKNEAHPYLKOLEGDLISLKEKNAHRIKTIISPTFHEINKLILVPLKSV 179
 DB 127 VLSKHFHFGKPRNRIGRLANGLVHDEKKAHRIINPAFHHEKIKGMPPVSTCC 186
 QY 180 TDVYKMSDKLSENGEVVDYEMFOLLTEVDVSRFAFGSSYDGAAYRLDAQMLCA 239
 DB 187 IEMITRWDSMPSGSEIDVWPEFCVLTGDISRAFGSNVYQGRIFELQELAERLI 246
 QY 240 EAPKVFIPGYRFFPTGNLKSRLDKRIKSLIKLIERROAIVGSEGECKEPPAAKL 299

DB 247 QSVOTIFIPGYWELPTKNNRRRAIDVEIRKLEIIIGREKDT-----KNRETNDDL 300
 QY 300 LGLMIOAK-----NTVODIVECKSPFAGKOTSNLTTTTLISMPERQAK 349
 DB 301 LGLLESNTROSGNASLGLTTEVDYIECKLFYRAGMETTIVLLTWLIVLSNHEWQER 360
 QY 350 ARDEVLRVGSRDVPTKDHVVKLTSLMINESLRLYPPIVATIRRAKSDVKLGKYPIC 409
 DB 361 AREEVLSHNG-RTTDPYDSLRLKTIITMILEVLRILYPPATFLRRYKEMELGKIKYPA 419
 QY 410 GTRELIPITIAVHHDAIWNADVNEFNPAFADGVPRAAKHVGFIPFGLSVRTICIGONLA 469
 DB 420 GVDLLPVTFIHDPINWCKDASEFNPQRFEDGINSATGHQAFFPFGMGPRICIGONF 479
 QY 470 ILQAKLTAVMIOREFFHLAPTYOAHPTVLMLYPOHGAFTFRRLT 516
 DB 480 LLEAKMTLCTILQRFSELSSTYTAHYVTTLHPHQAQIRLKL 526

RESULT 6

Q9FDZ1 PRELIMINARY; PRT; 523 AA.
 AC Q9FDZ1; 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE Putative cytochrome P450.
 GN P0688A04.10 OR P0006C01.25 OR RICE2.
 OS Oryza sativa (Rice), and
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoae; Oryza.
 ON NCBI_TaxId=4530; 35947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 clone: P0688A04.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 clone: P0006C01.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Imahishi H., Shimizu M., Ohkawa H.;
 RT "Cytochrome P450 in rice.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF002839; BAB19104.1; -;
 DR EMBL; AF002744; BAB19083.1; -;
 DR EMBL; AB047400; BAB85117.1; -;
 DR Gramene; G9FDZ1; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 523 AA; 59873 MW; E31AF21BC17A83CA CRC64;

Query Match 42.7%; Score 1165.5; DB 10; Length 523;
 Best Local Similarity 44.1%; Pred. No. 1e-87;
 Matches 231; Conservative 97; Mismatches 175; Indels 21; Gaps 7;

QY 5 SSSWTFPKVYLVSILSVIYKGMSLMWRPKRIEHSKQGINRPPHYFIGNVKELV 63

DB 8 ASQMTLAAAAAVALMLAVSTLEMAWMTERRALRAOGIRGNRYLFTGDVENV 67
QY 64 GMLLKASHMPFSS-HNILEPVLSPYHMRKLYGATFLWFGPTFRLTVADPDIREIFS 122
DB 68 RLNRERARKKLPGLGCHDILPRVLPMSKAVEHGRSPFTWGPTRWMSISPEIRVMS 127
QY 123 -KSEFYKNEAHPVLYQLEBGDLSLKERKAHRRKISPTFMENLKLVLYVLSKVD 181
DB 128 NKEGHYKPKPRLGKLL-ASGVSEYEGEKMAKHRIINPAFHEKIKEMLPVSNCTE 186
QY 182 MYDKMSDKLSENGEVEVDYEMFOLLTEDEVISRTAFGSSYEDGRAVFLQAQMLLQAA 241
DB 187 MYTRWNSWISBMSSEVDWPEFQNLTDGVLSKIPGSSYEGRRIFQLQESARLLOA 246
QY 242 FQKVPIDGYRFFPTRGNLKSRKLDKIRKSLKLTERRONAIDEGECKEPAKDLG 301
DB 247 FRTIFPGYWFPLTKNNRRLREIEREVSLLKLGIGKERAIKOG-----ETSGDLG 300
QY 302 LMIQAK-----NTVODIVECKSFPPAGKQTSNLTWTITLISHMPWOKAR 351
DB 301 LIVESNMRESNGKALGMTTDEIECKLFYAGMETTSVLLTTLVLSHMPWQERAR 360
QY 352 DEVLRYCGSDVPTKDHVVKLTLSMILNESLRYPPVATIRAKSPVKLGKXICGT 411
DB 361 EYLVLHFG-KITPDVDSLSRKIVMILYEVRLVLPVPLTRRYKEMELGKIKYPAV 419
QY 412 ELIPIAVHDOAIWGNVDNENPAPADGVPRAKAPVGFIPGLGVTCTICQNTAIL 471
DB 420 TMLPILFTHDPIDIGKAGSEFNPGFADGISNAKTQTSFFPGMGFRICIQNALL 479
QY 472 QAKTLAVMIQRPFTPLAPTYOHAFTVLMILYPOGAPITPRL 515
DB 480 EAKMAICTTIQRPSELSPTHTAFTVITLHPQAGQIKLKI 523

RESULT 7

Q9ATU3 PRELIMINARY; PRT; 525 AA.
AC Q9ATU3;
AT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Putative cytochrome P450.
OS Lolium rigidum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaceae; Lolium.
OX NCBI_TaxID=89674;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SLR 31;
RA Fischer T.C., Klattig J.T., Gierl A.
RT "A general cloning strategy of divergent plant cytochrome P450 genes
and its application in Lolium rigidum and Oculum basilicum."
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF321868; AAK38092.1; -
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PROSITE; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 525 AA; 59998 MW; DDFC8B96DFFP060C CRC64;

Query Match 42.3%; Score 1154.5; DB 10; Length 525;
Best Local Similarity 43.9%; Pred. No. 8.5e-87;
Matches 232; Conservative 97; Mismatches 175; Indels 25; Gaps 9;

DB 8 MIREASPMISLAGAAAAALLMLAAWYE--MAMTRRLRLRAQAGLACTQYRLFTGD 64
QY 59 VKELVGMMLKASHMPFSS-HNILEPVLSPYHMRKLYGATFLWFGPTFRLTVADPDIL 117
DB 65 VERNRLNRERARKKLPGLGCHDILQRVQPMNSVIXENKGRSPFTWGPTRWMSISPEIRV 124
QY 118 REIFS-KSEFYKNEAHPVLYQLEBGDLSLKERKAHRRKISPTFMENLKLVLYVVL 176
DB 125 REVLNKKGHYKQKSSRLGKLL-ANGIANHOGKMAKHRIINPAFHEKIKEMLPVPA 183
QY 177 KSVTMYDKMSDKLSENGEVEVDYEMFOLLTEDEVISRTAFGSSYEDGRAVFLQAQML 236
DB 184 TCCEEMITRWNMSNTESSSIDIMPEQNLTDGVISRTAFGSSYEGMKIFQLQESLAE 243
QY 237 LCAEAPQKVPIDGYRFFPTRGNLKSRKLDKIRKSLKLTERRONAIDEGECKEPA 296
DB 244 RLIMAFCOTIFPGYWFPLTKNNKMRALDCEIR-TIIRVIRKKOKAIXN-----GA 297
QY 297 KDLGLMIQAK-----NTVODIVECKSFPPAGKQTSNLTWTITLISHMPW 346
DB 298 DDLGLLESNMRESNGKADLGMTTDEIECKLFYAGMETTSVLLTTLVLSHMPW 357
QY 347 QAKADEVLRYCGSDVPTKDHVVKLTLSMILNESLRYPPVATIRAKSPVKLGKXIK 406
DB 358 QKKADEVLYHFG-KITPDVDSLSRKIVMILYEVRLVLPVPLTRRYKAMELGKIK 416
QY 407 IECGTEELIPIAVHDOAIWGNVDNENPAPADGVPRAKAPVGFIPGLGVTCTICQ 466
DB 417 YPAGVNLMLPILFTHDPIDIGKAGSEFNPGFADGISNAKHGSPFGMGFRICIQ 476
QY 467 NLAIIOAKTLAVMIQRPFTPLAPTYOHAFTVLMILYPOGAPITPRL 515
DB 477 NFALIEAKMALSTIIQHPSELSPTHTAFTVITLHPQAGQIKMKI 525

RESULT 8

Q9ATU2 PRELIMINARY; PRT; 525 AA.
AC Q9ATU2;
AT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Putative cytochrome P450.
OS Lolium rigidum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaceae; Lolium.
OX NCBI_TaxID=89674;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SLR 31;
RA Fischer T.C., Klattig J.T., Gierl A.
RT "A general cloning strategy of divergent plant cytochrome P450 genes
and its application in Lolium rigidum and Oculum basilicum."
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF321869; AAK38093.1; -
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PROSITE; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 525 AA; 59631 MW; 2C3032186914DF49 CRC64;

Query Match 42.2%; Score 1151.5; DB 10; Length 525;
Best Local Similarity 43.7%; Pred. No. 1.5e-86;
Matches 229; Conservative 96; Mismatches 178; Indels 21; Gaps 7;

QY 4 EESSNPIKPVLYSLVLYVVKGMSLIMPRKRIEHEFSQGRGPPHYFIQNVELY 63
DB 11 EASPNMLCAAAAVVWLA-AWILEWAMWTPRRGRALLEQKGRIRLFTGDVENV 69

QY 64 GMLKASSHPMPF-SHNILPRVLSFYHHMKRTYCATLWFGPFRLLTVADPDLLREIFS 122
 Db 70 RANKEARSKPLPLGSHDIIIPVQPMISNAIKENGLSTFWGPEPRVTLIDPEVRILIS 129
 QY 123 -KSEFYENKAEHPLVQLBGGDGLSLKGEKMAHRRKIISPFFHEMNLKLVPPVYLKSVTD 181
 Db 130 NKGHYGKRSSRFGLL-ANGLVNHQGEKMAKRRILNPAFHEKIKRMLPVSACSEE 188
 QY 182 MYDKMSDKLSENGEVEVDVYEWQILTEDIYSRTAFSGSYEDGAVFRLQAQMLLCAEA 241
 Db 189 MTRWENSMSSQGVSEVDVWPEFQNLTDVISRFAFGSSYQEGTKIFOLQGEQERLMOA 248
 QY 242 FOKVPIPGVFPFPTGNLKSRLDKERKSLKLIERRONALDGECEKEPAKDLG 301
 Db 249 FQTLFPGVFLPTKNRRMRERIDREICTILRGILIEKDRAIKSGEA-----SSDLLG 302
 QY 302 LMIQAK-----NTVODIVECKSPFFAGKOTTSNLLTWTTILLSHPEWQAKAR 351
 Db 303 LLBSNRBSNGKADLGSTEDLIECKLFYFAGMETTSVLLTWLILVSHPEWQEQAR 362
 QY 352 DEVLAVCGSDVPTKOHVVKLITSLMINESLRLYPPIVATIRAKSDVKLGKXIPCGT 411
 Db 363 KEVLHFFG-RTDPDEFNSRLKIVTWLYEVLRLPRAIPYTRRTYKAMELGITTPAGV 421
 QY 412 ELIPIIAVHDOAIMGNDVNEPNRPAFADGVPRAKHPVGFIPFGLGVRTICIGNTAIL 471
 Db 422 NMLPILFIHNDPNWKGDASEFNFORPADGISNAVKHPAFAFPFGGFRICIGNFALL 481
 QY 472 QAKTLTAVMIQRTFHLAPTQAPVLMILYPOGAIITRRRL 515
 Db 482 EAKMALSTILQRFSEFSPSYTHAPYVTLHPQGAIVLRKI 525

RESULT 9

Q9ATU1 PRELIMINARY; PRT; 525 AA.
 ID Q9ATU1
 AC Q9ATU1
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Putative cytochrome P450.
 OS Lolium rigidum.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Poaceae; Lolium.
 OC NCBI_TaxID=89674;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SR 31;
 RA Fischer T.C.; Klattig J.T.; Gierl A.;
 RT "A general cloning strategy of divergent plant cytochrome P450 genes
 and its application in Lolium rigidum and Octium basilicum";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF321870; AAK38094.1; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 525 AA; 59601 MW; B729F87AB214C1B9 CRC64;

Query Match 42.0%; Score 1147.5; DB 10; Length 525;
 Best Local Similarity 43.7%; Pred. No. 3.2e-86;
 Matches 229; Conservative 95; Mismatches 179; Indels 21; Gaps 7;
 4 ESSSWFIPKVLVLSVLSLVYVKGSLMWRPKRIEHSFGKIRGPPYHFGTGNVKELV 63
 11 EASPNLACAAAMAVLWLA-AWILEMAMTTPRLGRALERQIGKGTFRYLTGVDVENA 69

QY 64 GMLKASSHPMPF-SHNILPRVLSFYHHMKRTYCATLWFGPFRLLTVADPDLLREIFS 122
 Db 70 RANKEARSKPLPLGSHDIIIPVQPMISNAIKENGLSTFWGPEPRVTLIDPEVRILIS 129
 QY 123 -KSEFYENKAEHPLVQLBGGDGLSLKGEKMAHRRKIISPFFHEMNLKLVPPVYLKSVTD 181
 Db 130 NKGHYGKRSSRFGLL-ANGLVNHQGEKMAKRRILNPAFHEKIKRMLPVSACSEE 188
 QY 182 MYDKMSDKLSENGEVEVDVYEWQILTEDIYSRTAFSGSYEDGAVFRLQAQMLLCAEA 241
 Db 189 MTRWENSMSSQGVSEVDVWPEFQNLTDVISRFAFGSSYQEGTKIFOLQGEQERLMOA 248
 QY 242 FOKVPIPGVFPFPTGNLKSRLDKERKSLKLIERRONALDGECEKEPAKDLG 301
 Db 249 FQTLFPGVFLPTKNRRMRERIDREICTILRGILIEKDRAIKSGEA-----SSDLLG 302
 QY 302 LMIQAK-----NTVODIVECKSPFFAGKOTTSNLLTWTTILLSHPEWQAKAR 351
 Db 303 LLBSNRBSNGKADLGSTEDLIECKLFYFAGMETTSVLLTWLILVSHPEWQEQAR 362
 QY 352 DEVLAVCGSDVPTKOHVVKLITSLMINESLRLYPPIVATIRAKSDVKLGKXIPCGT 411
 Db 363 KEVLHFFG-RTDPDEFNSRLKIVTWLYEVLRLPRAIPYTRRTYKAMELGITTPAGV 421
 QY 412 ELIPIIAVHDOAIMGNDVNEPNRPAFADGVPRAKHPVGFIPFGLGVRTICIGNTAIL 471
 Db 422 NMLPILFIHNDPNWKGDASEFNFORPADGISNAVKHPAFAFPFGGFRICIGNFALL 481
 QY 472 QAKTLTAVMIQRTFHLAPTQAPVLMILYPOGAIITRRRL 515
 Db 482 EAKMALSTILQRFSEFSPSYTHAPYVTLHPQGAIVLRKI 525

RESULT 10

Q9ATU4 PRELIMINARY; PRT; 525 AA.
 ID Q9ATU4
 AC Q9ATU4
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Putative cytochrome P450.
 OS Lolium rigidum.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Poaceae; Lolium.
 OC NCBI_TaxID=89674;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SR 31;
 RA Fischer T.C.; Klattig J.T.; Gierl A.;
 RT "A general cloning strategy of divergent plant cytochrome P450 genes
 and its application in Lolium rigidum and Octium basilicum";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF321867; AAK38091.1; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 525 AA; 59894 MW; D202EC96A3946F14 CRC64;

Query Match 42.0%; Score 1145.5; DB 10; Length 525;
 Best Local Similarity 43.5%; Pred. No. 4.7e-86;
 Matches 230; Conservative 96; Mismatches 178; Indels 25; Gaps 8;
 1 MEESSSMFPKVLVLSVLSLV--IVKGSLLMWRPKRIEHSFGKIRGPPYHFGTGNV 58
 8 MURASPSVSLAGAAAMALMLAWIYE---WAMWTPRLRRLAQOGLRGTRYLFTGD 64
 59 VKELVGMKASSHPMPF-SHNILPRVLSFYHHMKRTYCATLWFGPFRLLTVADPDLL 117


```

Db      92 VLKSHGLKSPFWTSPKRPFFVITIRPDILAREILSKSGNFAKQTTAGIAKFAVVG-gvwtYEG 150
OY      150 EKMAHHRKTIISPTFHEMLTLTVLVYKLSVTMDVMDKSDKSENGEVEVDVYEMFOLLTE 205
Db      151 EKMAHRRILNPAHFOEKIRKMLPFLACCTKMTIRWVNSGSSGISELDWDEFQNLITG 210
OY      210 DIVISRTAPGSSYEDGRVFLPQAQOMLLCAEAFQKVFIFGYPFPFRGNLKSRLKDEIR 265
Db      211 DIVISRTAPGSSYEGGNRI FOLJEOEAKRVLKAFQRFIFIGYVYLIEJNNRRIREIDQIR 270
OY      270 KSLIKLERRRQNAIDEGEGECKEPAKDLGLIMIOAK-----NIVVQDIVESCK 315
Db      271 TILGIIYKPKDAVRNENSG-----DILGILVNSMRGSENEKEDVGMSEIDWIECK 325
OY      320 SEFPAGKQTSNLLTWTTTILLSMPEMQAARDEVLRVCGSDVPTKDHVVVKKLTSKIL 375
Db      325 LFVAGSETTSMLLTWTTTILLSMPEMGEQDEAEVWHNGF-RTTDPDHGLSRLLKIVMIL 385
OY      380 NESRLVPEPATIRAKSDVYKIGYKIPGCEELIPITAAVHDDAIGNVNFEFNPAF 435
Db      384 HEVRLRPVVFVFLQRTTHKEIEIGIKYBEGNVFLIPVLSIHDSISGQAIKFNPERF 440
OY      440 ADGVPRAAKHPVGFIPGIGVTRTCIQNTAIIQAKTLTAWIIOFTFHAFTYHAPTVL 495
Db      444 ANGVSKKATKCFQTAFFSFAMQPRICLGSSALLEAKVALATIIQSPFSELSYTHAPTV 505
OY      500 MLVYPOHGAFTTFRLL 515
Db      504 LTLOPQYGSPIKMKKL 519

RESULT 13
O9LUD2 PRELIMINARY; PRT; 515 AA.
ID ID O9LUD2
AC AC O9LUD2
DT DT 01-OCT-2000 (Tribble1.15, Created)
DT DT 01-OCT-2000 (Tribble1.15, Last sequence update)
DT DT 01-JUN-2003 (Tribble1.24, Last annotation update)
DE DE Cytochrome P450 (AT39J4620/MEI.12).
OS OS Arabidopsis thaliana (Mouse-ear cress).
OC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
CX CX NCBI_TaxId=3702;
RN RN [1]
RP RP SEQUENCE FROM N.A.
RC RC STRAIN=Columbia;
RC RC Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.,
RN RN Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RP RP [2]
RP RP SEQUENCE FROM N.A.
RC RC STRAIN=Columbia;
RX RX MEDLINE=20277480; PubMed=10819329;
RA RA Nakamura Y.;
RT RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT RT features of the regions of 4,504,864 bp covered by sixty P1 and PAC
RT RT clones.";
RL RL DNA Res. 7:131-135(2000).
RP RP [3]
RP RP SEQUENCE FROM N.A.
RA RA Cheuk R., Chen H., Kim C.-J., Koessens E., Meyers M.C., Banh J.,
RA RA Bowers L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA RA Ishida J., Jiang P.X., Jones T., Kamiya A., Katlin-Neumann G.,
RA RA Kawai J., Lam B., Lee J.M., Lin Y., Liu S.X., Miranda M., Narusaka M.,
RA RA Nguyen M., Onodera C.S., Palm C.J., Pam P.K., Quach H.T., Sakurai T.,
RA RA Saitou M., Setl M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA RA Ecker J.R.;
RT RT "Arabidopsis cDNA clones.";
RT RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CG CG -1. SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
RR RR EMBL; AB023038; BAB02394.1; -.
RR RR EMBL; AY052208; AA97679.1; -.

```

DR	GO:0004497; P:monoxygenase activity; IEA.
DR	GO:0006118; P:electron transport; IEA.
DR	InterPro: IPR001128; Cytochrome_P450.
DR	Pfam: PF00657; P450_1.
DR	PRINTS: PR00385; P450.
DR	PROSITE: PS00086; CYTOCHROME_P450; 1.
DR	Heme; Monooxygenase; Oxidoreductase.
DR	SEQUENCE 515 AA; 58643 MW; 9FAEFCE7686F55A1 CRC64;
Query Match	41.2%; Score 1124.5; DB 10; Length 515;
Best Local Similarity	40.3%; Pred. No. 2,5e+84;
Matches 209; Conservative 126; Mismatches 162; Indels 21; Gaps 6	
QY	12 KVLVLSVLSLVV-----KMSLIMPRKIEBHFSKQIGRPYPHFIGNKELV 63
DB	5 KIVAAVAAVVVVTVTVTWIMKGLNVAMLPRKNGEALYLRKQGLSGTFFLVGDIREA 64
QY	64 GMLLKASHPNPFSSHILPVLSPFHNHKKYGGATFLVWPGPTPLTVAPDILREIFSK 123
DB	65 SWVEGKSRPINLTDYTHRWMLIQYVHOGHTSYKMGRIASVTVTPHHIKVLN 124
QY	124 SEFYENEAHPLVKQLEGDGLLSKGEKWAHKKIISPTFMENLKLVVVLKSTVDNY 183
DB	125 VYDFPPVPHPIV-ELFATGVALYEGEWSYGRKIINPSFLEELTKIMIFAYESCEMI 183
QY	184 DKVSDLSNNGEV-EVDYVEMFOITLTVISRTFFGSGYEDGRVFLDAQOMLLCAEF 242
DB	184 SKMEKIVTQGSNNEIDWPRIIGDLSVIRTFSGSYEGSKIFLQEGGRVYLKL 243
QY	243 QKVFIEGYFPFTRGNLSRKLDKERKSLKILERRQNNADGEGECKEPAKOLLGL 302
DB	244 ELAFIEGKFFLPTKNNLRQKQINKEVSRRLREIIMKQRGMDTGEAK-----NDLLG 297
QY	303 MIQAKR---VAVODIVECKSPFPAGQTTNNLTITLISHNPEMAKADVELNVC 358
DB	298 LLESNGDGMGIEDVEBECKLFHNAQETIYALLVMTMLSHQNMQQAEBELKVI 357
QY	359 GSRDVTQDHYVVKLTLSMILNESIRLVPVATIRPAKSDVKLG-GYKIPCTELLIP 417
DB	358 GKNNKKNFALSRKLTMSMILNEVLRPLPGILLGRVVEKTKLGEEDMTLPGAQVPIV 417
QY	418 IAVHNGQALWGDVNEFPAPADVPAAAGHPVGFIPFGVGYTCIGUALIYAKLT 477
DB	418 IAVHNPRLMGDVEHNEFSPADISKATNQVSPFLPFGMGPRFGCGQFALMEKML 477
QY	478 AVMIQRFTHLAPTYQHPAPVLMILYPRQCAFIFRRL 515
DB	478 VLIQRFSPFLSPSYTHAPHTVLTLPDQGFALIFNML 515

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF446870; ALJ38603.1; -
 DR GO; GO:0004497; P:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 DR Heme; Monooxygenase; Oxidoreductase.
 KW SEQUENCE 515 AA; 58616 MW; 9P4AD6C768648CA1 CRC64;

Query Match 41.2%; Score 1124.5; DB 10; Length 515;
 Best Local Similarity 40.3%; Pred. No. 2.5e-84;
 Matches 209; Conservative 126; Mismatches 162; Indels 21; Gaps 6;

```

QY 12 KVLVLSVLSLVY-----KMSLLMWRPKIEHFKQIGRGPYHFFIGNVYELV 63
DB 5 KISAVAVAAVVVYVTVVWIMKGLNVAWLPKNEAYLKRQGLSGPFFLVGDYKRA 64
QY 64 GMMKASHPMPFESHNLPRVLSFYHMRKIYGATPLWFGPTFRLTVADPDILREIFSK 123
DB 65 SVWDEKSRPINTLDYTHRWVPLIQGVNDHGTYSMMGPIASVITKBEHIDVLR 124
QY 124 SEFYKNEAHPVYKQLEGDGLLSKGEKMAHRRKIISPTFMENIKLVVVLKSVTDVY 183
DB 125 YVDFKPPVHPV-ELFATGVALYEGEKMSKRRKIIPSPHLEKDKIMPAFYESGSEMI 183
QY 184 DKWSKLSNGEV-EVDVYEWFOITLDEVISRTFGSSYSDGRVAFRLQAQMLCAEAF 242
DB 184 SWEKLVLEQSSNIDVWPTLGDITSRTVSSRTVSSYIEGKRIFFLOEBOGRVYLKL 243
QY 243 QKVFYGRFFPFRGNLSKRLDKERKSLIKLERRONALIDEGECKEPAKDLIGL 302
DB 244 ELAFIPGRFLPTKSNLMRQINKEVKSRLREIMKQKMDTSEAK-----NDLIGI 297
QY 303 MIAQAK---VTVODIVECKSFPAAGQQTNSLITTTLLSHPEWQAKARDEVLRV 358
DB 298 LLESNGDHGSIEVVEECRLPHFAGQETAVLWMTLSSHQKQDAEELIKVI 357
QY 359 GSRDVPYKDVVVKLTSLINLESRLPYPIVATIRAKSDVYKLG-GYKIPCGTELLIPT 417
DB 358 GGNKPNPDALSLKTMNLNEVRLKLPGLIGRVEKETKIGEDMTLPGAQVYIPV 417
QY 418 IAVHDDQAIWQNDVNEFNPAPRAGQVPRPAKHVGFIFPGIGVATCGQNLALIQAKTL 477
DB 418 LMTVHDDPELMGEDVHEFPERPADGISKATKNQVSFLPFGWGRFCGQNFALMEAKVAL 477
QY 478 AVMIORFPHLAPYQHAPTVMLPYPOHGAIPTRRL 515
DB 478 VTILORFPELSPSTTHAPHTVLTILHPOGAPLIFHML 515

```

RESULT 15

08LGN4 PRELIMINARY; PRT; 528 AA.

ID 08LGN4; DB 08LGN4;
 AC 01-OCT-2002 (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE Cytochrome P450 monooxygenase CYP72A26.
 OS Zea mays subsp. mays (maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4578;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Inbred B73;
 RA Wang J., Schuler M.;
 RT "Characteristics and Transcriptional Regulation of Maize Cytochrome
 P450 Monooxygenases." to the EMBL/GenBank/DBJ databases.
 RL Submitted (DEC-2001)

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Inbred B73;
 RA Wang J., Reed R., Schuler M.;
 RT "Xenobiotics Regulate Maize Cytochrome P450 Monooxygenases at a Post-
 Transcriptional Level." to the EMBL/GenBank/DBJ databases.
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AY019866; AL60592.1; -
 DR GO; GO:0004497; P:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 DR Heme; Monooxygenase; Oxidoreductase.
 KW SEQUENCE 528 AA; 59972 MW; BAIEDBD207E6B60B CRC64;

Query Match 41.0%; Score 1119; DB 10; Length 528;
 Best Local Similarity 43.1%; Pred. No. 7.4e-84;
 Matches 214; Conservative 103; Mismatches 164; Indels 16; Gaps 7;

```

QY 28 MSLMWRPKIEHFKQIGRGPYHFFIGNVYELVGMKASHPMPF-SHNLPRVLS 86
DB 39 LRLMWRPKIEHFKQIGRGPYHFFIGNVYELVGMKASHPMPF-SHNLPRVLS 98
QY 87 FYHMRKIYGATPLWFGPTFRLTVADPDILREIFG-KSEFYKNEAHPVYKQLEGDGL 145
DB 99 FLHRLVLEHGTSNWFSPKVTIVDEPLAKVLSNKGHEFKLVKPSLKL-SGVYA 157
QY 146 SLKGEKMAHRRKIISPTFMENIKLVVVLKSVTDVWDM-SDKLSNGEVAVDYENF 204
DB 158 SHGEKMAHRRKIISPTFMENIKLVVVLKSVTDVWDM-SDKLSNGEVAVDYENF 217
QY 205 QITLEDVTSRTAFSSYEDGRVAFRLQAQMLCAEAFQKV-FIPGRFPPTGNLSRK 263
DB 218 QNLTGDVTSRTAFSSYEDGRVAFRLQAQMLCAEAFQKV-FIPGRFPPTGNLSRK 277
QY 264 LDKERKSLIKLERRONALIDEGECKEPAKDLIGLMIQAK-----NVTVDIVEEC 318
DB 278 NNREVESTILREILIKRIQAMERGGTK-----DMLGLLETNMRDDMGMTIEDVIEEC 331
QY 319 KSFFPAGKQTTNSLITTTLLSHPEWQAKARDEVLRVCGSDVPTKDVVVKLTLSM 378
DB 332 KVFYFAGMETTSVLLTMTVVLSSHPEWQAKARDEVLRVCGSDVPTKDVVVKLTLSM 391
QY 379 LNESRLKPYPIVATIRAKSDVYKGYKTCGTGLLPIAVHDDQAIWQNDVNEFNP 438
DB 392 LVEVRLTPPATSVVRQYKEMEVGVTPYAGVILFELPVLIIHDDPDWGDAREFRPDR 451
QY 439 PADGVPRAKHPVGFIPFGIGVATCGQNLALIQAKTLAVMIORFPHLAPYQHAPTV 498
DB 452 FSDGVSRASKDPGAFIPFGWGRFCGQNFALMEAKVALIIOREFRLAPSTHAPHT 511
QY 499 LMLYPOHGAIPTRRL 515
DB 512 VTILORFPELSPSTTHAPHT 528

```

Search completed: May 19, 2004, 10:27:20
 Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 10:24:45 ; Search time 23 Seconds
(without alignments)
1167.196 Million cell updates/sec

Title: US-09-992-901-2

Perfect score: 2730
Sequence: 1 MEESSSMFIPKVLVLSVIL.....LTPQHGAITFRRLNHD 520

Scoring table:

BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2730	100.0	520	4 US-09-527-073-2	Sequence 2, Appl1
2	482.5	17.7	489	4 US-09-852-067-4	Sequence 4, Appl1
3	479.5	17.6	524	4 US-09-976-594-533	Sequence 533, App
4	473	17.3	510	4 US-09-852-067-2	Sequence 2, Appl1
5	441.5	16.2	504	4 US-08-457-274A-25	Sequence 25, Appl1
6	441.5	16.2	504	5 PCT-US95-05758-25	Sequence 25, Appl1
7	418	15.3	503	4 US-09-144-367-2	Sequence 2, Appl1
8	404	14.8	576	3 US-08-948-564-16	Sequence 16, Appl1
9	386	14.1	503	4 US-09-583-447A-2	Sequence 2, Appl1
10	377.5	13.8	504	4 US-09-583-447A-4	Sequence 4, Appl1
11	367	13.4	557	4 US-09-518-386B-1	Sequence 1, Appl1
12	367	13.4	557	4 US-09-518-386B-3	Sequence 3, Appl1
13	347	12.7	502	4 US-09-904-615-69	Sequence 69, Appl1
14	344.5	12.6	507	4 US-08-457-274A-23	Sequence 23, Appl1
15	344.5	12.6	507	5 PCT-US95-05758-23	Sequence 23, Appl1
16	335.5	12.3	511	3 US-08-991-677-4	Sequence 4, Appl1
17	331.5	12.1	507	1 US-08-457-274A-22	Sequence 22, Appl1
18	331.5	12.1	507	5 PCT-US95-05758-22	Sequence 22, Appl1
19	329.5	12.1	496	3 US-09-292-768-66	Sequence 66, Appl1
20	329.5	12.1	517	4 US-09-302-620B-100	Sequence 100, App
21	329.5	12.1	517	4 US-09-911-781-32	Sequence 32, Appl1
22	328.5	12.0	496	3 US-09-292-768-64	Sequence 64, Appl1
23	327.5	12.0	496	3 US-08-881-784-1	Sequence 1, Appl1
24	327.5	12.0	496	3 US-09-292-768-2	Sequence 2, Appl1
25	327.5	12.0	496	3 US-09-127-339-6	Sequence 6, Appl1
26	323	11.8	520	4 US-09-564-808-4	Sequence 4, Appl1
27	323	11.8	540	4 US-09-302-620B-98	Sequence 98, Appl1

28	320.5	11.7	497	3 US-08-724-466B-4	Sequence 4, Appl1
29	320.5	11.7	497	4 US-08-882-164D-4	Sequence 4, Appl1
30	318	11.6	508	4 US-09-126-420A-25	Sequence 25, Appl1
31	318	11.6	540	4 US-09-302-620B-99	Sequence 99, Appl1
32	316.5	11.6	420	4 US-09-583-447A-6	Sequence 6, Appl1
33	316	11.6	476	1 US-08-313-075A-30	Sequence 30, Appl1
34	315	11.5	523	4 US-09-302-620B-95	Sequence 95, Appl1
35	314.5	11.5	497	4 US-08-882-164D-32	Sequence 32, Appl1
36	314.5	11.5	509	3 US-08-948-564-18	Sequence 18, Appl1
37	314.5	11.5	517	4 US-09-302-620B-101	Sequence 101, App
38	314	11.5	199	4 US-09-351-229-2	Sequence 2, Appl1
39	312.5	11.4	508	3 US-08-991-677-2	Sequence 2, Appl1
40	311	11.4	520	4 US-09-564-808-2	Sequence 2, Appl1
41	308	11.3	529	4 US-09-615-192A-405	Sequence 405, App
42	305.5	11.2	513	3 US-08-948-564-2	Sequence 2, Appl1
43	305	11.2	509	4 US-09-499-302A-8	Sequence 8, Appl1
44	305	11.2	520	2 US-09-091-432-2	Sequence 2, Appl1
45	305	11.2	520	4 US-09-387-663-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-527-073-2
; Sequence 2, Application US/09527073
; Patent No. 6534313
; GENERAL INFORMATION:
; APPLICANT: Michael M. Neff
; APPLICANT: Joanne Chory
; TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
; TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING
; FILE REFERENCE: SALKINS.024A
; CURRENT FILING DATE: US/09/527,073
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: US 60/124570
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: US 60/170,931
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: US 60/172,932
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-527-073-2
Query Match 100.0%; Score 2730; DB 4; Length 520;
Best Local Similarity 100.0%; Pred. No. 4.1e-278;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MEESSSMFIPKVLVLSVILSVYKMSILMRPKIIEHFEKQIRGPFYFGVXK 60
1 MEESSSMFIPKVLVLSVILSVYKMSILMRPKIIEHFEKQIRGPFYFGVXK 60
61 ELVGMMLKASHHPFSGHILPRLSPFYHMKIKYATFLVWGPFTFLVADPDILREI 120
61 ELVGMMLKASHHPFSGHILPRLSPFYHMKIKYATFLVWGPFTFLVADPDILREI 120
61 ELVGMMLKASHHPFSGHILPRLSPFYHMKIKYATFLVWGPFTFLVADPDILREI 120
61 ELVGMMLKASHHPFSGHILPRLSPFYHMKIKYATFLVWGPFTFLVADPDILREI 120
121 FSKSEFYKNEAPLVQLEGDGLSLKGEKMAHREKTIPTFHEMLKILVVVYKSVT 180
121 FSKSEFYKNEAPLVQLEGDGLSLKGEKMAHREKTIPTFHEMLKILVVVYKSVT 180
121 FSKSEFYKNEAPLVQLEGDGLSLKGEKMAHREKTIPTFHEMLKILVVVYKSVT 180
121 FSKSEFYKNEAPLVQLEGDGLSLKGEKMAHREKTIPTFHEMLKILVVVYKSVT 180
161 DMVDKMSKSENEVEVDVYEMEQILTEVYISRTAFSGSYEDGRAVFRLOAQOMLCAE 240
161 DMVDKMSKSENEVEVDVYEMEQILTEVYISRTAFSGSYEDGRAVFRLOAQOMLCAE 240
161 DMVDKMSKSENEVEVDVYEMEQILTEVYISRTAFSGSYEDGRAVFRLOAQOMLCAE 240
161 DMVDKMSKSENEVEVDVYEMEQILTEVYISRTAFSGSYEDGRAVFRLOAQOMLCAE 240
241 AFQKVFIFQYFFPTFRGLSKRKLDKEIRKSLKTIERRRQNAIDGEGECKEPAAKDL 300
241 AFQKVFIFQYFFPTFRGLSKRKLDKEIRKSLKTIERRRQNAIDGEGECKEPAAKDL 300
241 AFQKVFIFQYFFPTFRGLSKRKLDKEIRKSLKTIERRRQNAIDGEGECKEPAAKDL 300
```

```

QY 301 GIMIOAKVTVODIVEECKSFEPAGKOTSNLTWTITLISMPENQAKRDEVLVCCS 360
DB 301 GIMIOAKVTVODIVEECKSFEPAGKOTSNLTWTITLISMPENQAKRDEVLVCCS 360
QY 361 RDVPTKDEHVVLKLTLSMLINESLRYPEIVATIRAKSDVYLGSKYKIPCTELLIPILAV 420
DB 361 RDVPTKDEHVVLKLTLSMLINESLRYPEIVATIRAKSDVYLGSKYKIPCTELLIPILAV 420
QY 421 HHDAIWKNDVNEFPAFADGVPRAAHNPVGFIPGLGVRTICGNLALLOAKTLAVW 480
DB 421 HHDAIWKNDVNEFPAFADGVPRAAHNPVGFIPGLGVRTICGNLALLOAKTLAVW 480
QY 481 IORFTFHLPYQAHAPYVLMILYPOHGAPITFRRLTNHED 520
DB 481 IORFTFHLPYQAHAPYVLMILYPOHGAPITFRRLTNHED 520

```

RESULT 2

```

US-09-852-067-4
; Sequence 4, Application US/09852067
; Patent No. 6531297
; GENERAL INFORMATION:
; APPLICANT: MERKTOV, Genady et al
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,
; FILE REFERENCE: CLO00897-CIP
; CURRENT APPLICATION NUMBER: US/09/852,067
; CURRENT FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Human
US-09-852-067-4

```

```

Query Match 17.7%; Score 482.5; DB 4; Length 489;
Best Local Similarity 27.1%; Pred. No. 8.8e-42;
Matches 141; Conservative 103; Mismatches 215; Indels 61; Gaps 15;

```

```

QY 20 LSLVIYKGMSTL---WMPRKIEEHFSKQGRGPPYHFGNVEKGMMLKXSHMPF 76
DB 7 LLLILKKAQQLYLRQWLLRALQQ-----FPCPFHMLSHSRE-----F 46
QY 77 SHNLLPVLSEYHMRKTYGATFLVWF-GPTFRITVADPDILIREIFSKSEFEYKNEAHP 135
DB 47 QND---QELERIQKWEKFPGACPMWLSGNKARLLVDPYKILGRSD-PRAPRYVKL 102
QY 136 VKQLGGDGLSLKGRKMAHNRKIISPTHEMNLKLVPLVYLKSTDMVDKMSKLSNGE 195
DB 103 MPMWIGYGLLLDGGTWFQHRMLTPAHYDILKPYGLMWDVQIMLDMEQLISDDSS 162
QY 196 VEVADVEMFQILTEDEVISRTAF---GSSYEDGRAVFRLOAQOML--LCAEAFQKVF--- 246
DB 163 LB--IFQVSLMTLDTIKKCAFSYQGSVDLRNHSYIQAINDMLNVFRANVFIQSD 220
QY 247 -----IPGYRFPPTGNLKSRLDKETIKSLKLTERRRQNALDGESECKEPAKADLG 301
DB 221 FLVRLSPGRGLFHRACQLAHEHTDVIDO-----QRKAOQOGSELEKVRKRRLDFLD 273
QY 302 LMIQAK-----NVTQDIVEECKSFEPAGKOTSNLTWTITLISMPENQAKRDEVLV 356
DB 274 VILFAKMGSSLSQDLRAVDTFMFEGHDTTASGVSMFYALATPHRQHRCHREHIOG 333
QY 357 VCGSRDVPYKDVVLKLTLSMLINESLRYPEIVATIRAKSDVKL-GYKIPCTELLIP 415
DB 334 LIGDGAISLWHEHIDMPYTTMCIKKALRLYPVSVSRQLSKPTFPDGRSLPBGVLTFL 393
QY 416 PIIAVHHDAIWKNDVNEFPAFADGVPRAAKHPVGFIPGLGVRTICGNLALLOAKL 475
DB 394 SIYGLHYMKWQNP-EVDFPFRFA---PDGAYSHALPLPSGARBCIGQFAMRELVK 449

```

```

QY 476 TLAVMIQRTFTHLPYQAHAPYVLMILYPOHGAPITFRRL 515
DB 450 AVALTLIRFELLPDPIFRPIPIARVVLKSKNTHLRKL 489

```

RESULT 3

```

US-09-976-594-533
; Sequence 533, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROI
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 533
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte ID No. 6673549 6825202CD1
US-09-976-594-533

```

```

Query Match 17.6%; Score 479.5; DB 4; Length 524;
Best Local Similarity 28.4%; Pred. No. 2e-41;
Matches 154; Conservative 83; Mismatches 207; Indels 99; Gaps 21;

```

```

QY 1 MEESSSNFIPKVLVLSVLSLVIYKGMSTL---WMPRKIEEHFSKQGRGPPYHFP 55
DB 1 MPQLSLWLGGLGVPAAFPWILLLVGGSWILARVLAW-----TYTF 41
QY 56 IGVNKELVGMMLKXSHMPFSNHLPRVLSFYHMRKI-----YGATF 99
DB 42 YDNRRL-----QCFQF-----PKQWPFHGQGLVTFEEGKTTLTQLVTTYPQF 88
QY 100 LVMEGPTFRITV-ADPDILIREIFSKSEFEYKNEA--HPLVKQLGGDGLSLKGRKMAHR 156
DB 89 KLMIGPITPILLCHPILIRPISASAAPKXMIYFGFLKPMJGGLLSGDDKXSR 148
QY 157 KIISPTHEMNLKLVPLVYLKSTDMVDKMSDKLSNGEVEVDVYEMFQILTEDEVISRTA 216
DB 149 RMLTPAFHFNILKEVYMKIFKXSVIMHDKW-QLASBGSARLDMFHHISLMTLDSQKCV 207
QY 217 FGSSYEDQ-----RAVFRLOA-----QQMLCAEAFQKVFIPGYRFPPTGNLKS 261
DB 208 F--SFENCGKSEYIYALIELSAPYERKNOQILHTDLYLVLTGQGRF-----RBA 259
QY 262 RKLDEIKRSLKLTERRRQNALDGESECKEPAK---DLGLMT-----QAKNTVOD 313
DB 260 CHLVHDFDAVITQ--ERRRLLPTCGIDDFLKNKXSKTIDFIDVLLISXDEDEKLSDD 317
QY 314 IVECKSFEPAGKOTSNLTWTITLISMPENQAKRDEVLVCCSRDVPK---DHVY 370
DB 318 IRAEDTWFEGHDTTASGLSWLYHLAKRPETQSOOROVQGLADRE-PIEIMDDLA 376
QY 371 KLTLSMLINESLRYPEIVATIRAKSDVKL-GYKIPCTELLIPILAVHHDAIWK 429
DB 377 QLPFTWICKSLHLPVPVPIISRCTQDPEVLDPGRVIFPGIYCLINIIGHNPTVW-P 435
QY 430 DVNEFPAFADGVPRAAKHPVGFIPGLGVRTICGNLALLOAKTLAVMIQRTFTHLA 489
DB 436 DEVDYDFRR-DOENIKERSPLAIFPSAGPRNCIGQAFMAEMKVLTALT--LIFRIL 492
QY 490 PTY 492
DB 493 PTH 495

```

```

RESULT 4
US-09-852-067-2
; Sequence 2, Application US/09852067
; Patent No. 6531297
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: C0000897-CIP
; CURRENT APPLICATION NUMBER: US/09/852,067
; CURRENT FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Human
US-09-852-067-2

Query Match      17.3%; Score 473; DB 4; Length 510;
Best Local Similarity 25.7%; Pred. No. 9,4e-41;
Matches 138; Conservative 111; Mismatches 221; Indels 68; Gaps 19;

QY      8 WFIKRVLVSVILSVYKNSLIMWPKIEHFSKQIRGPPYHFGNVELVGMML 67
DB      10 WAPRYLAFFVFCALGLQAIKYLRRQRLRD---LRPPAPPTWFLGHQKFI----- 61
QY      68 KASHMPFESHNLPRVLSFHWKRYGATFLVPGPTR--LVADPLIIEIRSKSE 125
DB      62 --ODNNWKELEITEK-----YPPAPFPWICP-FOAFCTYDDPYAKTILSRID 107
QY      126 --FYEKNEAHPYKQLEGDGLSLKGEKWAHHRKIIISPTFHEINIKLVVYLVKSVTDM 182
DB      108 PKSRYLQKFSPLD---GKGLALDPRKMFQHRRLITPFGFHNILKAYIEVWASHVKRM 163
QY      183 VDKSDDKLSNGEVEVDYEWFOILTEDVISRTAFG-----SYED-GRVFRP---- 230
DB      164 LDKW-EKICSTQDSVVEYEHINMSLDIIMKCAFSEKTCQNTSTHDPYAKIIFELSKI 222
QY      221 ---QAQWMLCAEAFQKFIPIGYRFPPTGKLSKRLDEIRKSLKLIERRQNAIDSE 287
DB      223 IFRHRYSLYHSDIIFGLSPGGRF-----QKSRVLANQYTTIIDERKKSLOAGV 273
QY      288 GEECKEPAK--DLGLMTIOAKNVTVQ---DIVECKSFPPAGKOTTSNLTWTITLL 340
DB      274 KOD-NTPKRYQODELDYLSAKDESGSPFDIDVHSEVSFTLLAGHDTLAASISWILYCL 332
QY      341 SMHPWOKARDEVLRVCGSDVPTKDHVVYKLTSLINLESRLVZPPIVATIRRAKSDV 400
DB      333 ALNPHQRCREBRGCIIGDGSSTTWOLGEMSTTMCIKETCRILPAVVISIRDSKYL 392
QY      401 KL-GGYKIPCGTELLIPPIAVHND-QAIWGNVNEFPNAPFADGVPRAAHPVGFIPFGL 458
DB      393 TFPDQCTLPAGITVYVLSIMGLHNPAAVWKNP-KVFDPLRFSQ-ENSDQHPPAVAYPFS 450
QY      459 GVRTTCIGNALIQAOKTLAVMIQRTFHILAPYQAHPTVLMELLYPHOGAPIFRRYLT 516
DB      451 GSRNCTGQEFMTEIKVTILILHFRVTPDPTPLTFPHNFIKPKNGMYILAKXLS 508

```

```

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,274A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1304
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; STRADEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Rat
; STRAIN:
; DEVELOPMENTAL STAGE: Adult
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
US-08-457-274A-25

Query Match      16.2%; Score 441.5; DB 1; Length 504;
Best Local Similarity 28.1%; Pred. No. 1.9e-37;
Matches 141; Conservative 95; Mismatches 185; Indels 81; Gaps 19;

QY      18 VILSLVYKNSLIMWPKIEHFSKQIRGPPYHFGNVELVGMMLKASHMPFS 77
DB      13 VLAIVLV-----LLYGGTHTGGLFKQGIPIGB-----KPLPF- 46
QY      78 HNLIPVLSFYHH-WR-----KIYGATPIVWFG--PTRLTVADBDLIREIFSK---SE 125
DB      47 ---FGVLANYMGLWKEVDCHKKYKIGLFDGQMLPAIT--DTEMKNVAVKSCFSV 101
QY      126 FYEKNEAHPYKQLEGDGLSLKGEKWAHHRKIIISPTFHEINIKLVVYLVKSVTDMVK 185
DB      102 FTRNRDGFV--GIMGKAVSVAODEWKRRLALSTPFSGRLEKMFPII-EGYGLIVK 158
QY      186 WSDTLSENGEVEVDYEWFOILTEDVISRTAFG---SYEDGAIVRLQAQWMLCAEAF 242
DB      159 YLKQAEATGK-FVTMKVFGAYSMVDYTSISFGVNVDSLNPPDPE-VEKTKLLEFDF 216
QY      243 QKVFIPIGYRFP--PTGKLSKRLDEIRKSLKLIERRQNAIDGEECKEPAKDL 299
DB      217 DPLSLVYLPFLPLPIYEMINICFPKDSIEFPKKVYMKERLDS-----VQKRVDP 271
QY      300 LGIMIOAKN-----VYQDIVECKSFPPAGKOTTSNLTWTITLLSMP 344
DB      272 LQIMMNAHNSKDESHSTALSDEWETASII-----FIFGYPTSTISFVLHSLATHP 326
QY      345 EMOAKARDEVLRVCGSDVPTKDHVVYKLTSLINLESRLVZPPIVATIRRAKSDVKG 404
DB      327 DTQKLGEBEIDRALPNAPPTIYDVMEMLDVLNMTLRLIYIGNRLEBYCKDVEING 386
QY      405 YKIPCGTELLIPPIAVHNDQAIWGNVNEFPNAPFADGVPRAAHPVGFIPFGLGVRTCT 464

```


Mon May 24 08:10:10 2004

us-09-992-901-2.rat

Page 4

Db 387 VEMPKGSVVMIPSYALHBDPQHW-DEPEEFPERFSK-ENKGSIDPYVILPFGNGPNCI 444
QY 465 GQNTALLQAKLTLAVMIQRTF 486
Db 445 GMRFLMNMKALTITVQLNFSF 466

RESULT 6
PCT-US95-05758-25
Sequence 25, Application PC/TUS9505758
GENERAL INFORMATION:
APPLICANT: Cornell Research Foundation, Inc.
TITLE OF INVENTION: Cytochrome P450Ipr Gene and Its
TITLE OF INVENTION: Uses
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05758
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael J.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1600
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rat
STRAIN: Unknown
DEVELOPMENTAL STAGE: Adult
POSITION IN GENOME:
CHROMOSOME/SEGMENT: Unknown
PCT-US95-05758-25

Query Match 16.2%; Score 441.5; DB 5; Length 504;
Best Local Similarity 28.1%; Pred. No. 1.9e-37;
Matches 141; Conservative 95; Mismatches 185; Indels 81; Gaps 19;

QY 18 VILSVLYVKMSLMMRPKRIEHEFSKQIGRPYHFIQNGKELVGMMLKASSHPMPS 77
Db 13 VLLAVLV-----LLYGFGRTHGLFKKQIGPQ-----KPLDF- 46
QY 78 HNLIPRLSYFHH-KR-----KIYGATFLVWG--PFRILTVAOPDLIREIFSK---SE 125
Db 47 ---FGVTLNLYWGLMRFVDECHKKYKIGLFGDGMPLFAIT--DTEMIKAVLYECPSV 101
QY 126 FYRNEAHPVKQLEBDDGLISLKGEMAHHRKIIPTFIMENLKLIVVYLKSVTDMDK 185
Db 102 FTTRRFPGPV--GIMGAVSVADDEMKRYRALISFTFSGLKEMFPII-EQIGDILVK 158
QY 186 WSKLSENGEVEVDVEMFQILTEDVISRTAFG---SSYEDGRAVFLQAQGMMLCAEAF 242

Db 159 YLKDAEAGK-PVTMKVFAGAYSMVDVITSFGVNVDSLNNPQDP-VEKTKLLEPDPF 216
QY 243 QKVFIPGYR---PFRGNLKSRLDKEIRKSLILKIERRONAIDGEGECKEPAKOL 299
Db 217 DELPISVYLFPFLPIYEMINICMFPPKDSIEFFKKFYRNKERILDS-----YQKRVDF 271
QY 300 LGMTQAQN-----VTYQDIVECKSFYFKQKOTTSWLTWTITLISMP 344
Db 272 LQLMNANHDSKDXESHSTALSDMEITAOSII-----FIFAGYEPTSTSLFVLSLATHP 326
QY 345 EMOAKRPEVLRVCGSRDVPFKDHHVVLKTLMSLINESLRYPIVATIRAKSDVLAG 404
Db 327 DQCKLGEIRBALPNKAPPTDYVMEMEYLDVMTETLLIYIGNRLEKCKKDVING 386
QY 405 YKIPGTELLIPILAVHHDAQIWGNDVNEFNPAFADGVRPAKHVGFIPGLGYRTCI 464
Db 387 VEMPKGSVVMIPSYALHBDPQHW-DEPEEFPERFSK-ENKGSIDPYVILPFGNGPNCI 444
QY 465 GQNTALLQAKLTLAVMIQRTF 486
Db 445 GMRFLMNMKALTITVQLNFSF 466

RESULT 7
US-09-144-367-2
Sequence 2, Application US/09144367
Patent No. 6432639
GENERAL INFORMATION:
APPLICANT: Iacheter, Jay
TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4
FILE REFERENCE: SEQ-12P
CURRENT FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/038,612
PRIOR FILING DATE: 1997-09-10
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 503
TYPE: PRT
ORGANISM: H. sapiens
US-09-144-367-2

Query Match 15.3%; Score 418; DB 4; Length 503;
Best Local Similarity 27.3%; Pred. No. 5.6e-35;
Matches 139; Conservative 92; Mismatches 185; Indels 94; Gaps 18;

QY 14 LVLSVLYSVYKMSLMMRPKRIEHE-FSKQIGRPYHFIQNGKELVGMMLKASS 71
Db 13 LLLAVSVLYLYG-----THSHGLFKKQIGPQ----- 42
QY 72 HMPFSNHLIPRLSYFHH-----HWKTYGATFLVWGPFRLTVADPDLIREIFSK 123
Db 43 -PLPLGNILS-----YKGFQCFDMECHKKYGVWGVYDQCVLAITPDMIKIVLVK 96
QY 124 ---SEFYENEAHPVKQLEBDDGLIS-----LKSEKAHHRKIIPTFIMENLKLIVPV 175
Db 97 ECVSVFTNRPRPGV-----GFWKSAISIAEBEWRGLSLSPFTFSGLKEMFPII 149
QY 176 LKSVTMDVMDKDSNGEVEVDVEMFQILTEDVISRTAFG---SSYEDGRAVFLQA 232
Db 150 AQ-YEDVLYVRNRRALAGTK-PYTLKQVFGAYSMVDVITSFGVNVDSLNNPQDP-VEN 206
QY 233 QQMLCAEAFQKVFIPGYR---PFRGNLKSRLDKEIRKSLILKIERRONAIDBGE 289
Db 207 TKKLRFPLDPPFLSITVFPLPILEVLNLCVFPREVNTFLKSVRMKESRL----- 261
QY 290 ECKEPAKQDLGIMTQAQN-----VTYQDIVECKSFYFKQKOTTSWLTWTITL 340
Db 262 EDTQKHVDVFLQMLDSQNSKATESHKAISDLDELVAQSIIIFAGYETTSVLSFIMVEL 321

QY 341 SMHEPWOAKARDEVLKRVGSRDVPVTKDHVVKLTSLMINESLALPYVATIRAKSDV 400
 Db 322 AHFDVQCKLQSEIDALPDKAPPTVTVLQWETLDMVNETLPLIARLERVCKOV 381
 QY 401 KLGGYKIPCGTELLPIIAVHDDAIWGNVNEFNPAFA---DGVPRAKHPVGFTPF 456
 Db 382 EINGNEIPKGVVMTIPSYALHNDPKYM-TSEKFLPERFSKXNDKI----DPYITPF 435
 QY 457 GLGVFTICIGONLAILIQAULTLAVMOTRETF 486
 Db 436 GSGPRNCIAGRFALMMKALIRVQNFSF 465

RESULT 8

US-08-948-564-16
 ; Sequence 16, Application US/08948564
 ; Patent No. 6121512

GENERAL INFORMATION:

APPLICANT: Simmszy, Balazs
 APPLICANT: Dewey, Ralph E.
 APPLICANT: Corbin, Frederick T.
 TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
 TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Virginia C. Bennett
 STREET: PO Box 37428
 CITY: Raleigh
 STATE: No. 6121512th Carolina
 COUNTRY: USA
 ZIP: 27627

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/948,564
 FILING DATE:
 CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Bennett, Virginia C.
 REGISTRATION NUMBER: 37,092
 REFERENCE/DOCKET NUMBER: 5051-409
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-854-1401
 TELEFAX: 919-854-1401
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 576 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

US-08-948-564-16
 ; MOLECULE TYPE: Protein

Query Match 14.8%; Score 404; DB 3; Length 576;
 Best Local Similarity 26.8%; Pred. No. 2.1e-33;
 Matches 133; Conservative 93; Mismatches 201; Indels 70; Gaps 19;

QY 56 IGVNVELVGMUKASS-HMPFSH---NILPRVLSF-YHMRKIYGAFTLVWFGPTFL 109
 Db 60 LQNAENLLTDLGSGISGSMPIAGAVSDILGRFLPSLYDWLEHGAAYKLAFGPKAV 119
 QY 110 TVADDLIREIFSKSEF-YEKNEAHPVYKOLEGDLSTLGEKMAHRRKISPTPMENYL 168
 Db 120 VVSDIVARHILREAFSYDKGLADLIEPIWGLIPALDITWQRRVYIAPAFNSYL 179
 QY 169 KLLVPLVLSVTDWYDKSKLSENG---EVEVDYEMFOITLTDVTSRTAFSSYED 223
 Db 180 EAMVRIPTGSERTILKENKLLBEGYDGPDSITLDEAFSSIALDIIGLVF--NDVF 237
 QY 224 G-----RAVFLQAGQMLCAEAPQKVFIPGRFPFTRGNL-KSRLLD--KEIRK 270

Db 238 GSVTKESPVIAKY-----GTLPEAHRSTFIPIYWKILARWIVPROKFDLKYINT 292
 QY 271 SLKTIERRRQNAIDSGECKECP---AAKD--LIGLMIAQKNVTYQD--IVECKSEFF 323
 Db 293 CLDGLRNNAKESRQETDVEKLLQORDYLNKDSALRFLVDRGADVDRQRLDMLMLI 352
 QY 324 AGKOTSNLTWTITLLSMHEPWOAKARDEVLKRVGSRDVPVTKDHVVKLTSLMINESL 383
 Db 353 AGHETIAAVTVAVFLAQNPSKMKKAQAVDVLGT--GRPTFESKELQYIRIIVYEAAL 411
 QY 384 RLYPPVATIRRA-KSDVYKGGYK-----IPCTELLPIIAVHDDAIWGNVNEFN 435
 Db 412 RLYPQPLIRRLSKSDVLPQGHKQKQGYALPAQTDVAFISVYMLHRSPPYFMDRP--DDFE 470
 QY 436 PARF-----ADG-----VPRAKHPVGFIPIGCVFTICIGONLAILIQAOK 474
 Db 471 PERFLVQNNKEIEGAGLDPBSRFGALYPNEVISDFALFPGGPRKCVGDFALWEST 530
 QY 475 LTLAVMIQRTFHLAFT 491
 Db 531 VALTMLQNFVDELKGT 547

RESULT 9

US-09-583-447A-2
 ; Sequence 2, Application US/09583447A
 ; Patent No. 6645745

GENERAL INFORMATION:

APPLICANT: WOJNOMSKI, Leszek
 APPLICANT: GELNER, Klaus
 APPLICANT: EISELT, Regina
 TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A
 TITLE OF INVENTION: (CYP3A) GENE FAMILY: CYP3A
 FILE REFERENCE: 310115.401
 CURRENT APPLICATION NUMBER: US/09/583,447A
 CURRENT FILING DATE: 2000-05-30
 NUMBER OF SEQ ID NOS: 45
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 2
 LENGTH: 503
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-583-447A-2

Query Match 14.1%; Score 386; DB 4; Length 503;
 Best Local Similarity 25.9%; Pred. No. 1.3e-31;
 Matches 128; Conservative 93; Mismatches 202; Indels 72; Gaps 16;

QY 22 LVIVKMSLLMWRPKIEHSKQIRPPPHFPIGNVKELVGMUKAASHMPFSNLT 81
 Db 13 VIVATSLVLLVYGTSHSKLFRKGLIPPT-----PLPFLGTL 51
 QY 82 PVLSPYHMRKI---YGATFLVWFGPTFRLTVADPLIREIFSK---SEFYKNEAHL 135
 Db 52 FYLRGLMNPDRCEKNEYGEMGLYEGQOPMLVIMDPDKITVLKECVSYFTNQMPGPM 111
 QY 136 VKQBEDGLS-----LKGKMAHRRKISPTPMENYLKLVVPLVLSVTDWYDKSKL 190
 Db 112 -----GFLKALSAEDEEMKRIKTLISALFISVAFKEMWPII--SQCGDMVLSRQE 163
 QY 191 SENGVEVDVYEMFOITLTDVTSRTAFG---SSYEDGAVFRLQAQOMLCAEFQKVF 247
 Db 164 AENSK-SIMKDFGAYTWDVITGTLFGVNDLSLNPDDP-LQNMKKLKL-DLDPFL 220
 QY 248 PGYRFP---PFRGNLKRKLDKXIRKSLKTLIRBRQNMIDGSECKECPAKDLGLM 303
 Db 221 LLSLPLPLTVPEALNIGLPKQDVTHPLKSIERKMSRL-----KDKQHRVDFQOM 275
 QY 304 IQAKN-----TVQDIVECKSPFPAGKQTTNLTWTITLLSMHEPWOAKARDEV 354
 Db 276 IDSQNKETKSHKALSDLELVAOSIIIFAAVDTISTTLPIVWEALATHDPVQCKQBEI 335
 QY 355 LRVGSRDVPVTKDHVVKLTSLMINESLRLYPPVATIRAKSDVYKGGYKIPCGTELL 414

```

Db      336 DAVLPNKAPVTVDALVQMEYLDVWVNETLRLFPVSVTVTRCKDKDLEINVFIPKGLAVM 395
Qy      415 IPIIVVHDDQIMWGDVNEFPNAPFADGVPRAAHPVG---FIPFGIGVATCIGONLAIL 471
Db      396 VPIYALHDDPKIV-TEPEKFCPEBSRS---KKNDSDIDLYRIYIPFGAGPANCIGMFAL 450
Qy      472 QAKLTAVMIQRTF 486
Db      451 NIKLAVIRALQNFSS 465

```

```

RESULT 10
US-09-583-447A-4
; Sequence 4, Application US/09583447A
; Patent No. 6645745
; GENERAL INFORMATION:
; APPLICANT: NOJONOSKI, Leszek
; APPLICANT: GELNER, Klaus
; APPLICANT: EISELT, Regina
; TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A
; FILE REFERENCE: 310.115.401
; CURRENT APPLICATION NUMBER: US/09/583,447A
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-583-447A-4

```

```

Query Match      13.8%; Score 377.5; DB 4; Length 504;
Best Local Similarity 25.8%; Pred. No. 1e-30;
Matches 128; Conservative 94; Mismatches 201; Indels 73; Gaps 17;

Qy      22 LVIVKMSLLMWRPKIEEHFSKQIRGPYHFFIGNVELVGMMLKASHPMPESHNL 81
Db      13 VLVATSLVLYITGTSKILFKLGIPT-----PLEPLIGTL 51
Qy      82 PVLSPFYHWKRI---YGATFLWFGPTFLVADPDLIREIFSK---SEFEKNEANPL 135
Db      52 FYLRGLMNFDECKEKGEMWGLYEGQCEMLVIMPDMIKTVLVEKCVSFTNQPLGPM 111
Qy      136 VKOLEGGLS-----LKGEKMAHHRKIISPTFHEMNLKLVAVVTKSVTDVDEKMDKL 190
Db      112 -----GFLKSAISPADEDEMKRITLSPAFTSVKFKEMVPII-SQCGMLVNSLROE 163
Qy      191 SENGVEVDVYEFQILTEDVISRIFAG---SSYEDGRAVFLQAOQMLCAEAFQYFI 247
Db      164 AENSK-SINKDFGAVTMDVITGLFGVLDISLNNPODF-LKMMKLLKLD-FLDPFL 220
Qy      248 PGRFF-----PTKGLKSRKLDKIRKSLKLIERRRQNALDEGECEKPPAKOLGIM 303
Db      221 LLISLFPFLTPVEALNIGLFPDOVTHFLKNSIERKESRL-----KOKQKRVDFPQOM 275
Qy      304 IOAKN-----VTQDIVECKSEFFAGKOTISNLLTWTTILSNMPEWQAVARDEV 354
Db      276 IDGNSKETSKHSLDELVAQSIIIFAAVDTISLTLPIFYELATHPVOQKLOEBI 335
Qy      355 LRVGSRDVPVKDHYVTKLTLMSILNESLALYPPIYATIRAKSDVKLGKYLPGSTEL 414
Db      336 DAVLPNKAPVTVDALVQMEYLDVWVNETLRLFPVSVTVTRCKDKDLEINVFIPKGLAVM 395
Qy      415 IPIIVVHDDQIMWGDVNEFPNAPFADGVPRAAHPVG---FIPFGIGVATCIGONLAIL 470
Db      396 VPIYALHDDPKIV-TEPEKFCPEBSRS---KKNDSDIDLYRIYIPFGAGPANCIGMFAL 450
Qy      471 LQAKLTAVMIQRTF 486
Db      451 NIKLAVIRALQNFSS 465

```

```

RESULT 11
US-09-518-386B-1
; Sequence 1, Application US/09518386B
; Patent No. 6365386
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
; CURRENT APPLICATION NUMBER: US/09/518,386B
; CURRENT FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: EP 99104668.1
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: EP 00101666.6
; PRIOR FILING DATE: 2000-02-01
; SOFTWARE: Patentin Ver. 2.1
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 1
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
; FEATURE:
; NAME/KEY: TRANSIT
; LOCATION: (1)..(26)
US-09-518-386B-1

```

```

Query Match      13.4%; Score 367; DB 4; Length 557;
Best Local Similarity 22.2%; Pred. No. 1.5e-29;
Matches 128; Conservative 111; Mismatches 203; Indels 134; Gaps 21;

Qy      13 VLVSVILSVIVKMSL---LWMPRKIEEHFSKQIRGPY-HFFIGNVELVGMML 67
Db      4 LVILGALGLAFAFWSAIFSLYLAPR-----SSLVNLQSPNHTFTGFGFDLS--- 56
Qy      68 KASHPMPESHNLRLVLSFYHWKRIYIGATFL---WFGPTFLVADPDLIREIFSKS 124
Db      57 -----ARTEBHAKTRKYGSLTRAGLNGAPVINST--DPVEVHWKMA 100
Qy      125 EPEYK-NEAHPVKOLEGGLSLKGEKMAHHRKIISPTFHEMNLKLVAVVTKSVTDV 183
Db      101 YDIPKGMARVRLATDGVYTAGEAHKEHRRIMISLQAQVKSVPFLFEKGMELV 160
Qy      184 DKMSDKLSN-----GEVE-----VDYEWFPQILTEDVISRIFAGSSYE----- 222
Db      161 DKMMDAARKMAVAGESAGEKKATRLTEGVADVKGWRATLDVWALAGFYKSDSLQNK 220
Qy      223 -----DGRAVFLQAOQMLCAEAFQYV---FIPGYRFP-----TRGNLK 260
Db      221 TNELYAFGLDGRAP-----FLDSFKAIMDFVYFPTMKRRHEIPULQGLAV 270
Qy      261 SRKLDKIRKSLKLIERRRQNALDEGECE---KEPAKOLLGIMQIA-----KN 308
Db      271 SRRVG-----IELMEQKQAVLSASDQAVRKQOVQGRDLSLVANANLAALPESQK 323
Qy      309 VTYQDIVECKSEFFAGKOTISNLLTWTTILSNMPEWQAVARDEVLRVCGSRDVPVKD 368
Db      324 LSDEEVLQISNLRPRGYTSSTVLTWPHRHSBKAVQDKLREICQI--DTMPFLIDE 381
Qy      369 VVKLTLSMILNESLRLYPPIYATIRAKSDVK-----LGQYKIPCGREL 413
Db      382 LNLPLYLEAFVYESLRLDPPSPYANRECKEDFIPLAEPVIGRDSVINEVRITKTMV 441
Qy      414 IPIIVVHDDQIMWGDVNEFPNAPFADGVPRAAK-----HPVGFIPFGIGVATCIGQ 466
Db      442 MLPFNINSKFIYGDABEFPERLBYVTDLSIASIPYGHQASFIS---GPRACGM 498
Qy      467 NLAIIQAKLTAVMIQRTF--LAFTYQHAFTVL 499
Db      499 RFAVAEWKAFLFVTLARVQPEPILSHPEYEHITLII 534

```

Mon May 24 08:10:10 2004

us-09-992-901-2.ral

Page 7

RESULT 12
US-09-518-386B-3
; Sequence 3, Application US/09518386B
; Patent No. 636386
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, TATSUO
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
; CURRENT APPLICATION NUMBER: US/09/518,386B
; PRIOR FILING DATE: 2000-03-03
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: EP 00101666.6
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
US-09-518-386B-3

Query Match 13.4%; Score 367; DB 4; Length 557;
Best Local Similarity 22.2%; Pred. No. 1,5e-29;
Matches 128; Conservative 111; Mismatches 203; Indels 134; Gaps 21;

QY 13 VLVSLVSLVIVKMSL---LWRRPKIEHRSKQIRPPY-HFETGVKELVGMVL 67
DB 4 LVLLTALGALPASMALFESLYLAPRR---SLVNLQSNPTFTNPFIDLS--- 56
QY 68 KASHPMPSPSHNLPVLSFYHMRKIKGATFL--VFGEPTFLVADBDLREIFSKS 124
DB 57 -----ARTGEHAKYREKXGSLRFAGIAGAVLNST--DKVFNHMKKA 100
QY 125 EPEYK-NEAPLVKQLEQDGLSLGEMKRAHRIKISTFMEMLKLVVVLKSTLDV 183
DB 101 YDPKPMARVLRIATGCVVTAEGEAKHRRKIMTFSLSAQVSKMVEIFLEKMEVL 160
QY 184 DKMSDKLSEN-----GEVY-----VDVYEWFOILTEDVISRTAFSSYE----- 222
DB 161 DKMEDPAEKMAVGSAGEKATRLTEGVVDXDMGRATLDVMAIAGFDYKSDLSQNK 220
QY 223 -----DGAAPVLOAQMLCAEFQKY--FIPGYRFP-----TGKNIK 260
DB 221 TNELVYAVGLTDPAP-----TLDSFKAIMDFVYEFTRKRHEIPLTQGLAV 270
QY 261 SRKLDKIRKSLKLIERRQNAIDGEGEC--KEPAKDLGLMIQA-----XN 308
DB 271 SRRVG-----IEMLEKQKQAVLGSASDAVDKDDVQGRDILSLVIRANIANLPESQK 323
QY 309 VWVQDIVECKSPFPAGKQTTSNLTWTITLLSHWPEKQAKAPREVRYVGSQSDVTKH 368
DB 324 LDBEVLQISNLRFGVETSSVLTMMFRLSEDKAVQDKLRBEICQI--DTDMPLDIB 381
QY 369 VVRLKTLISMLNESRLPYPIVATIRAKSDVK-----LGGYKIPQGTETL 413
DB 382 LNLPLFLAFLVKSRLDPPSPYANREGLKXEDFPLAEPVIGDSGVINEVITGTVMV 441
QY 414 LPIIIVVHNDQAIWGDVNEFNPAPFADGVPRAK-----HPVGGIPFGIGVRCICQ 466
DB 442 MLEPLNINSKFTYGDSEEFERKMLDVTDSIANSIEAPYGHQASFTS--GPRACFGR 498
QY 467 NLAIIQAKULTAVMIQRTFH--LAPTYQAPTVL 499
DB 499 RFAVEMKAFLEVTLRVQFEBIISHPEYEHITLI 534

RESULT 13
US-09-904-615-69
; Sequence 69, Application US/09904615
; Patent No. 6566325

GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 69
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (502)
; OTHER INFORMATION: Xaa equals stop translation
US-09-904-615-69

Query Match 12.7%; Score 347; DB 4; Length 502;
Best Local Similarity 26.0%; Pred. No. 1,6e-27;
Matches 129; Conservative 87; Mismatches 196; Indels 84; Gaps 22;

QY 49 GPHHPFIGNVKEIVGMMLKASHPMPSPSHNLPVLSFYHMRKIKGATFLVWFGPTFR 108
DB 42 GPHLPFTIGNISLA-----ASSP-----LPHV--YMKQSVYEITSLSLGLST 86
QY 109 LTVADPDLIRE-IFSKSEPEYKNEAHPV-VKQLESDGLSLK-GEKMAHRIKISPTFH 164
DB 87 VLVNGVDVVECELVHQSIEIFADRPCLEFMKQTKWGLNRRYGRGWDHRLAVNFRY 146
QY 165 -----MENKLVVVLKSTLDVMDVKMSDKLSENGSEVVDVYEWFOILTEDVISRT--AF 217
DB 147 FGVGQSSPESKILIEFTKFNDAIETKGR-----PDKQKINAVNSNTNLIIIF 196
QY 218 GS--SYEDG--RAVFLQAOQMLCAEFQKYFIPGYRFPFTRGKSRKLDKIRKS-- 271
DB 197 GERFTYEDTDFQMIELFSENVELAASA--SVFL--YNAFPWIGILPFGKHQQLFRMAAV 252
QY 272 -----LKLTERRR-----QNAIDGSEBEKKEPAADLGLMIOAKX-----VTVQDI 314
DB 253 VYDFLSRLKASVNEKPOLQHFVDAYLD-----NDQKNDSSSTFSKENTL 300
QY 315 VEECKSPFAGKQTTSNLTWTITLLSHWPEKQAKAPREVRYVGSQSDVTKHVVKLKT 374
DB 301 IFVSGELIAGTETTNVLRMAILFMALYPNIGQVQVEIDLIMGPNKSPWDXCKMPY 360
QY 375 LSMILNESRLPYPIVATIRAKS--DVLCGYKIPCGEMLLPIIIVVHNDQAIWGDVNE 433
DB 361 TEAVLHVLAFCQIVPLGIFAHISDQAVRGISIKGTIVTINLXSVHFDKXV-RDPEV 419
QY 434 ENRAPADGVPRAPKAPVGFIPFGIAGVTCIGQNTAIIQAKULTAVMIQRTFH--LA 489
DB 420 FHPERFDSSGYPAPKE-ALVDFSLGRHCHGEHARWEMELFTTALDORFHLHPHELV 478
QY 490 PTYQHAFTVLMILYPO 505
DB 479 PDKR--FRIGMTLQPO 492

RESULT 14
US-08-457-274A-23
; Sequence 23, Application US/08457274A
; Patent No. 5734086
; GENERAL INFORMATION:
; APPLICANT: Scott, Jeffrey G.
; APPLICANT: Tomita, Takashi
; TITLE OF INVENTION: Cytochrome P4501pr Gene and Its Uses

Mon May 24 08:10:10 2004

us-09-992-901-2.ra1

Page 8

NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSES: Nixon, Hargrave, Devans & Doyle
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,274A
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1304
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: drosophila melanogaster
STRAIN:
DEVELOPMENTAL STAGE: Adult
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
US-08-457-274A-23

Query Match 12.6%; Score 344.5; DB 1; Length 507;
Best Local Similarity 23.2%; Pred. No. 3.1e-27;
Matches 122; Conservative 116; Mismatches 211; Indels 77; Gaps 19;
Query 19 ILSLIVKMSLIMWPKIEBFSKQIRGPPYHFFIGNVELVGMMLKASHPMPFSH 78
Db 3 VLIYLLAISLILAYLYHNFNMYNRRGLPHDAPRLYGN--WVGRKRVWMDFFY-- 57
Query 79 NILPRVLSFYHHRKIYGATPL-VWGPFFRLTVADPLIREFSK--SEFEKNEAH-- 133
Db 58 -----DYNNKRR-SGFPFVGFFLHKPAAFIVDTOLAKNLIKDFSNADRGQFHNG 109
Query 134 ---PLVKQLEGGLSLKGEKVAHHRKIISTFHEKNTKLVVPLKSVTDMVKSPDKL 190
Db 110 RDPPLTQH-----LFLNDGKKWKDKRQLTFTSGKMKFPPVIAKVSSEFVAVITTEOV 164
Query 191 --SENGEVVDVYEMFQILTEVYISTAFSGSYEDGR---AFRLQAQOMLCAEAPKV 245
Db 165 PAQONAV-LEIEMARFTTDTIGTCRFGICNTLRTPVSDFRMGOKVTDNRHGL 223
Query 246 FIDYRFFPTRGKLDKEIRKSLKLIIE---RRQNAIDGEECKEPAADL 300
Db 224 TMVFSEPKLASRLRBMMEVDVHGFPMKLVNDTIALRERNF-----KRNDFM 272
Query 301 GLMTQAK-----NWTYQDVECKSEFPFGKQTSNLLTWTITLLSMRE 345
Db 273 NLIIEIKOKGSFTLNDGEVIEGMDIGLAADVFFVYAGFTSSSTSYCLYELAQND 332
Query 346 WQAKADEVLAVCGSRDVPDKHVVAKLT-LSMLNESLRLVPPVATIRRAKSDVKLG 404
Db 333 IQORLNEIGTVLEBEGQUTYSIRKMTYLVINGVISETLRLTVLPHLERKALNDYVVG 392

Query 405 YK--IPCGTELLIPIIAVHHDDQIMGNDVNEFNPAFADGVR--AAKPVGFIPEGG 459
Db 393 HEKVIENGTQVITIPACAYHRDEDLVNP-ETDPERFS--PEKVAARSVWMLPFGDG 448
Query 460 VRTICGNLAILQAKTLTAVMIQFTFLAPT-----YQHAFTYML 501
Db 449 PRNCIGRRFGOMCARIGIACIISFRVSVCDTEIPLKSPMSIVL 494
RESULT 15
PCT-US95-05758-23
Sequence 23, Application PC/TUS9505758
GENERAL INFORMATION:
APPLICANT: Cornell Research Foundation, Inc.
TITLE OF INVENTION: Cytochrome P4501pr Gene and Its
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSES: Nixon, Hargrave, Devans & Doyle
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05758
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1304
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: drosophila melanogaster
STRAIN:
DEVELOPMENTAL STAGE: Adult
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
PCT-US95-05758-23
Query Match 12.6%; Score 344.5; DB 5; Length 507;
Best Local Similarity 23.2%; Pred. No. 3.1e-27;
Matches 122; Conservative 116; Mismatches 211; Indels 77; Gaps 19;
Query 19 ILSLIVKMSLIMWPKIEBFSKQIRGPPYHFFIGNVELVGMMLKASHPMPFSH 78
Db 3 VLIYLLAISLILAYLYHNFNMYNRRGLPHDAPRLYGN--WVGRKRVWMDFFY-- 57
Query 79 NILPRVLSFYHHRKIYGATPL-VWGPFFRLTVADPLIREFSK--SEFEKNEAH-- 133
Db 58 -----DYNNKRR-SGFPFVGFFLHKPAAFIVDTOLAKNLIKDFSNADRGQFHNG 109
Query 134 ---PLVKQLEGGLSLKGEKVAHHRKIISTFHEKNTKLVVPLKSVTDMVKSPDKL 190
Db 110 RDPPLTQH-----LFLNDGKKWKDKRQLTFTSGKMKFPPVIAKVSSEFVAVITTEOV 164

```

QY 191 --SENBEVDVTEMFQIITEDVISTRTAFSSYEDGR---AVFRLQAQOMLCAEAFQKV 245
Db 165 PAQNGAV--LEIKELMARFTTDVIGTCRFIEGNTLRFPSDFTMGQKVFDMRGKLL 223
QY 246 FIDGYRFPFTRGNLKSRLDKETRSKSLKLE----RRONAIDGEBECPAPAKDIL 300
Db 224 TWFVSPFKLASRLRMRMBEDVHQPFMELVNDTIALRENF-----KRNDFM 272
QY 301 GLMIQAK-----NVTQDIVECKSFFFAQQTNSLLTWTTLISMPH 345
Db 273 NLLIELKQKGSFTLNDGEVIEGMDIGELAQVFWYVAGFETSSSTWSYCLYELAQND 332
QY 346 WQAKARDEVLRVCGSRDVPFKDHVVKLT--LSMILNESRLYPEIVATIRAKSDVKLG 404
Db 333 IODRLNVEIQTLEBQGGQLTYESITAMTYLNOVISETLRLTYLVPHERKALNDYVVP 392
QY 405 YK--IPCGTELLIPLIAVHDOAINQNDVNEENPARFADGVPR--AAKHVGFIPFGIG 459
Db 393 HEKLVIEKGTQVILIPACAYRDEDLXNP-ETPDPERFS--PEKVAABESVEMLPFGDG 448
QY 460 VRTICIGONTAIIQAKTLTAVMIQRTFHLAPT---YQHAFTVLM 501
Db 449 PRNCIGRFGOMQARIGLAQIISRFRVSVDTEIPLKXSPMSIVL 494

```

Search completed: May 19, 2004, 10:28:27
 Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 10:27:26 ; Search time 49 Seconds

(without alignments)
2952.979 Million cell updates/sec

Title: US-09-992-901-2

Perfect score: 2730

Sequence: 1 MEESSSWFIPKVLVSVIL.....LYPGGAPITFRRLTHED 520

Scoring table:

BIOSUM62

Searched: 1145568 seqs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries
Published Applications AA:
1: /cgn2_6/ptodaca/2/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodaca/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodaca/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodaca/2/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodaca/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodaca/2/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodaca/2/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodaca/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodaca/2/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodaca/2/pubppa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodaca/2/pubppa/US09C_NEW_PUB.pep:*
12: /cgn2_6/ptodaca/2/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodaca/2/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodaca/2/pubppa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodaca/2/pubppa/US10C_NEW_PUB.pep:*
16: /cgn2_6/ptodaca/2/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodaca/2/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodaca/2/pubppa/US60_PUBCOMB.pep:*Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2730	100.0	520	US-09-992-901-2	Sequence 2, Appl1
2	1386.5	50.8	533	US-10-424-599-202258	Sequence 202258,
3	1165.5	42.7	547	US-10-425-114-61218	Sequence 61218, A
4	1165.5	42.7	560	US-10-425-114-59349	Sequence 59349, A
5	1165.5	42.7	562	US-10-425-114-59350	Sequence 59350, A
6	1155.5	42.3	547	US-10-425-114-65616	Sequence 65616, A
7	1138.5	41.2	524	US-10-424-599-183328	Sequence 183328,
8	1124	41.7	508	US-10-424-599-195559	Sequence 195559,
9	1114	40.8	533	US-10-424-599-265344	Sequence 265344,
10	1103.5	40.4	537	US-10-425-114-38180	Sequence 38180, A
11	1102	37.5	524	US-10-424-599-217648	Sequence 217648,
12	1024	30.5	520	US-10-424-599-228819	Sequence 228819,
13	989	36.2	458	US-10-425-114-50377	Sequence 50377, A
14	973.5	35.7	532	US-10-425-114-58587	Sequence 58587, A
15	967	35.4	540	US-10-425-114-61220	Sequence 61220, A

16	967	35.4	552	US-10-425-114-64681	Sequence 64681, A
17	962	35.2	541	US-10-425-114-61219	Sequence 61219, A
18	930.5	34.1	527	US-10-425-114-67055	Sequence 67055, A
19	894	32.7	455	US-10-424-599-247367	Sequence 247367,
20	858	31.4	388	US-10-424-599-197466	Sequence 197466,
21	664.5	24.3	288	US-10-425-114-48579	Sequence 48579, A
22	660.5	24.2	546	US-10-424-599-219019	Sequence 219019,
23	543.5	19.9	505	US-09-910-689-307	Sequence 307, App
24	543.5	19.9	505	US-10-058-2700A-2	Sequence 2, Appl1
25	543.5	19.9	505	US-10-296-606-10	Sequence 10, Appl
26	543.5	19.9	505	US-10-010-742-307	Sequence 307, App
27	541	19.8	286	US-10-424-599-28055	Sequence 28055,
28	540.5	19.8	505	US-10-219-535-30	Sequence 30, Appl
29	540.5	19.8	505	US-10-232-230-30	Sequence 30, Appl
30	540.5	19.8	505	US-10-232-224-30	Sequence 30, Appl
31	540.5	19.8	505	US-10-067-668-2	Sequence 2, Appl1
32	540.5	19.8	505	US-10-227-884-30	Sequence 30, Appl
33	540.5	19.8	505	US-10-230-163-30	Sequence 30, Appl
34	540.5	19.8	505	US-10-230-338-30	Sequence 30, Appl
35	540.5	19.8	505	US-10-218-631-30	Sequence 30, Appl
36	540.5	19.8	505	US-10-230-414-30	Sequence 30, Appl
37	540.5	19.8	505	US-10-216-159A-30	Sequence 30, Appl
38	540.5	19.8	505	US-10-218-849-30	Sequence 30, Appl
39	540.5	19.8	505	US-10-227-873-30	Sequence 30, Appl
40	540.5	19.8	505	US-10-227-883-30	Sequence 30, Appl
41	540.5	19.8	505	US-10-219-076-30	Sequence 30, Appl
42	540.5	19.8	505	US-10-230-434-30	Sequence 30, Appl
43	540.5	19.8	505	US-10-219-003-30	Sequence 30, Appl
44	540.5	19.8	505	US-10-219-075-30	Sequence 30, Appl
45	540.5	19.8	505	US-10-219-464-30	Sequence 30, Appl

ALIGNMENTS

```
RESULT 1
US-09-992-901-2
; Sequence 2, Application US/09992901
; Patent No. US20020073446A1
; GENERAL INFORMATION:
; APPLICANT: Neff, Michael M.
; TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
; FILE REFERENCE: SAKIKINS, 024DVI
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US/09/992, 901
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: US 60/124570
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: US 60/170, 931
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: US 60/172, 832
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-992-901-2
Query Match 100.0%; Score 2730; DB 9; Length 520;
Best local similarity 100.0%; Pred. No. 5e-266;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEESSSWFIPKVLVSVILVYVGMGLMKRPKRIEHSKQIRPPHFFGNK 60
1 MEESSSWFIPKVLVSVILVYVGMGLMKRPKRIEHSKQIRPPHFFGNK 60
61 ELVGMMLKASHMPFESHNTLPVLSFYHMKRIYATFVMPGPTFRLLVADPDILRI 120
```


Db 61 ELVGMMLKASSHPMPFSNHLPRVLSFYHNRKICYGATFLVWPGPTPLTVADPDILREI 120

QY 121 PFSKSEFYKNEAHPVVKOLEBGGILSLKGEKMAHHRKISPTFMENLKLIVPVVLSKV 180

Db 122 PFSKSEFYKNEAHPVVKOLEBGGILSLKGEKMAHHRKISPTFMENLKLIVPVVLSKV 180

QY 181 DMVDKMSDKLSENGEVEVDVYEMFQILTEDVISTAFGSSYEDGRAVFRLOAQOMMLCAE 240

Db 181 DMVDKMSDKLSENGEVEVDVYEMFQILTEDVISTAFGSSYEDGRAVFRLOAQOMMLCAE 240

QY 241 AFQKVFIPGYRFFPRGNLKSRLDKETIRKSLKILIRRRONALDGECECKEPAKDL 300

Db 241 AFQKVFIPGYRFFPRGNLKSRLDKETIRKSLKILIRRRONALDGECECKEPAKDL 300

QY 301 GLMIOAKNVTVDIVECKSPFFAGKQTTSNLITWTITLLSMHEWQAKARDEVLRVCGS 360

Db 301 GLMIOAKNVTVDIVECKSPFFAGKQTTSNLITWTITLLSMHEWQAKARDEVLRVCGS 360

QY 361 RVPFKRDHYVVKLTSLMINESELPRIVATIRAKSDVKGXKIPCGTELLIPITAV 420

Db 361 RVPFKRDHYVVKLTSLMINESELPRIVATIRAKSDVKGXKIPCGTELLIPITAV 420

QY 421 HHDOAIWGNVDVEFPAPAFADGVPRAAKHPGFIPIFGIGVATCIGOMLAILOAKTLTAV 480

Db 421 HHDOAIWGNVDVEFPAPAFADGVPRAAKHPGFIPIFGIGVATCIGOMLAILOAKTLTAV 480

QY 481 IORFTFHLPYQHPATVLMILYPOHGAPIFRRLTNHED 520

Db 481 IORFTFHLPYQHPATVLMILYPOHGAPIFRRLTNHED 520

RESULT 2

US-10-424-599-202258

/ Sequence 202258, Application US/10424599

/ Publication No. US20040031072A1

/ GENERAL INFORMATION:

/ APPLICANT: La Rosa Thomas J

/ APPLICANT: Kovalic David K

/ APPLICANT: Zhou Yihua

/ APPLICANT: Cao Yongwei

/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

/ FILE REFERENCE: 38-21(53223)B

/ CURRENT APPLICATION NUMBER: US/10/424.599

/ CURRENT FILING DATE: 2003-04-28

/ NUMBER OF SEQ ID NOS: 285684

/ SEQ ID NO 202258

/ LENGTH: 353

/ TYPE: PRT

/ ORGANISM: Glycine max

/ FEATURE:

/ NAME/KEY: unsure

/ LOCATION: (1)..(353)

/ OTHER INFORMATION: unsure at all Xaa locations

/ FEATURE:

/ OTHER INFORMATION: Clone ID: PAT_MFT3847_24664C.1.pep

US-10-424-599-202258

Query Match 50.8%; Score 1386.5; DB 12; Length 353;

Best Local Similarity 76.8%; Pred. No. 9.1e-131;

Matches 214; Conservative 29; Mismatches 41; Indels 13; Gaps 5;

QY 1 MEESSSWFIPRVLVLSLVIVKGSLLMWRPKIEHESKQIGRPPYHFIQNVK 60

Db 1 MEESSSWFIPRVLVLSLVIVKGSLLMWRPKIEHESKQIGRPPYHFIQNVK 60

QY 61 ELVGMMLKASSHPMPFSNHLPRVLSFYHNRKICYGATFLVWPGPTPLTVADPDILREI 120

Db 61 ELVGMMLKASSHPMPFSNHLPRVLSFYHNRKICYGATFLVWPGPTPLTVADPDILREI 120

QY 58 ELVGMMLKASSHPMPFSNHLPRVLSFYHNRKICYGATFLVWPGPTPLTVADPDILREI 117

Db 58 ELVGMMLKASSHPMPFSNHLPRVLSFYHNRKICYGATFLVWPGPTPLTVADPDILREI 117

QY 121 F-SKSEFYKNEAHPVVKOLEBGGILSLKGEKMAHHRKISPTFMENLKLIVPVVLSKV 179

Db 121 F-SKSEFYKNEAHPVVKOLEBGGILSLKGEKMAHHRKISPTFMENLKLIVPVVLSKV 179

QY 118 FSKSEFYKNEAHPVVKOLEBGGILSLKGEKMAHHRKISPTFMENLKLIVPVVLSKV 177

Db 118 FSKSEFYKNEAHPVVKOLEBGGILSLKGEKMAHHRKISPTFMENLKLIVPVVLSKV 177

QY 180 DMVDKMSDKLSENGEVEVDVYEMFQILTEDVISTAFGSSYEDGRAVFRLOAQOMMLCA 239

Db 178 VEMLEKMS-AMGEKGYEIVSEMFQSLTEDVITRTAFGSSYEDGRAVFRLOAQOMMLCA 236

QY 240 EAFQKVFIPGYRFFPRGNLKSRLDKETIRKSLKILIRRRONALDGECECKEPAKDL 299

Db 237 DAFQKVFIPGYRFFPRGNLKSRLDKETIRKSLKILIRRRONALDGECECKEPAKDL 295

QY 300 LGMLIOA-----KNTVODIVECKSPFFAGKQTTSNLITWTITLLSMHEWQAK 349

Db 296 LGMLIOASNNMNSNTVDDMVECKSPFFAGKQTTSNLITWTITLLSMHEWQAK 352

RESULT 3

US-10-425-114-61218

/ Sequence 61218, Application US/10425114

/ Publication No. US20040034888A1

/ GENERAL INFORMATION:

/ APPLICANT: Liu, Jindong

/ APPLICANT: Zhou, Yihua

/ APPLICANT: Kovalic, David K.

/ APPLICANT: Screen, Steven E

/ APPLICANT: Tabaska, Jack E

/ APPLICANT: Cao, Yongwei

/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

/ FILE REFERENCE: 38-21(5313)B

/ CURRENT APPLICATION NUMBER: US/10/425.114

/ CURRENT FILING DATE: 2003-04-28

/ NUMBER OF SEQ ID NOS: 73128

/ SEQ ID NO 61218

/ LENGTH: 547

/ TYPE: PRT

/ ORGANISM: Oryza sativa nipponbare

/ FEATURE:

/ OTHER INFORMATION: Clone ID: LIB3431-054-B3_F11.pep

US-10-425-114-61218

Query Match 42.7%; Score 1165.5; DB 12; Length 547;

Best Local Similarity 44.1%; Pred. No. 3.5e-108;

Matches 231; Conservative 97; Mismatches 175; Indels 21; Gaps 7;

QY 5 SSSWFIPRVLVLSLVIVKGSLLMWRPKIEHESKQIGRPPYHFIQNVK 63

Db 32 ASQWILMAAAVAALLMLAVSTLEWAMTPTLRALAQDIRGRVRLFGVDPEV 91

QY 64 GMLKASSHPMPFS-NHILPRVLSFYHNRKICYGATFLVWPGPTPLTVADPDILREI 122

Db 92 RLNRKARKPLPLGCHDILPRVLPFSKAVESKGSFTWGPPTVMSIDPSIRVNS 151

QY 123 -KSEFYKNEAHPVVKOLEBGGILSLKGEKMAHHRKISPTFMENLKLIVPVVLSKV 181

Db 152 NKFGHYGKPKRTSLGKLL-ASGVVSYGEGMAHRLINPAFHEKIKMLPVFSNCTE 210

QY 182 DMVDKMSDKLSENGEVEVDVYEMFQILTEDVISTAFGSSYEDGRAVFRLOAQOMMLCAE 241

Db 211 MYTRMENSMSIEGSEVDVWPEFQNLGVDISTAFGSSYEDGRAVFRLOAQOMMLCAE 270

QY 242 FQKVFIPGYRFFPRGNLKSRLDKETIRKSLKILIRRRONALDGECECKEPAKDL 301

Db 271 FRITFIPOYWLPLPKNNRRLREIERVSKLRGILGRERAING-----ETSGDILG 324

QY 302 LMIQAK-----NVTYQDIVECKSPFFAGKQTTSNLITWTITLLSMHEWQAK 351

Db 335 LIVESNREBSKALQMTDEILIECKLFYFAGMETYSVLLTWITLIVSMHEWQAK 384

QY 352 DEVRVCGSRDVPFKRDHYVVKLTSLMINESELPRIVATIRAKSDVKGXKIPCGT 411

Db 385 EELVHNG-RTPPYDLSRLKIVTMLIVLVLPVYVLRRTYKEMELGGIKTPAEV 443

QY 412 ELLIPITAVHHDOAIWGNVDVEFPAPAFADGVPRAAKHPGFIPIFGIGVATCIGOMLAI 471

Db 412 ELLIPITAVHHDOAIWGNVDVEFPAPAFADGVPRAAKHPGFIPIFGIGVATCIGOMLAI 471

Db 444 TLMPLFIHHDPIWKGDAEFGNPGFADGISNATKQTSFFPFMGPRICIGONFALL 503

Qy 472 QAKLTLLAVMTQRTFTHLAPTYQAHAPTYMLLYPOHGAFTTFRRL 515

Db 504 EAKKAICTIIQRFSEFELSPSYIHAPFTVITLHPHOGAQTILKXI 547

RESULT 4

US-10-425-114-59349

Sequence 59349, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jindong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(5313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 59349

LENGTH: 560

TYPE: PRT

ORGANISM: Oryza sativa nippondare

FEATURE:

OTHER INFORMATION: Clone ID: JC-OSLEI1B3474019G01_FLI.pep

US-10-425-114-59349

Query Match 42.7%; Score 1165.5; DB 12; Length 560;

Best Local Similarity 44.1%; Pred. No. 3.6e-108;

Matches 231; Conservative 97; Mismatches 175; Indels 21; Gaps 7;

Qy 5 SSSMFIPKVLVSLVSL-VIVKMSLMMRPRIEHSKQIRGPPYHFIQNKELY 63

Db 45 ASQMTLAAAAAVALMLAVSTLEWMTPRLEBALRAGQIRNRLTGDVENV 104

Qy 64 GMLKASHPMPSS-HNLPRLVSYHMRKIYGAFTLVFGPTFRITVADPLIREIFS 122

Db 105 RLNRKARKPLPGCHDIIIRVLPMFSKAVEHGKPSFTWPGTPRVMISDESIREVWS 164

Qy 123 -KSEFYKRNHAPLVKQLEBDGSLSKGKMAHHRKIISPTHEMLKLLVAVLSVTD 181

Db 165 NKFGHYGKPKPTRLGKLL-ASGVASYEGKMAHHRILNPAFHHEKIKMLVPSNCTE 223

Qy 182 MYDKSDKLSNGEVEVDYVEMFQILTEDVISRTAGSSYEDGRAVFLQAOQMLCAEA 241

Db 224 MYRWENMSIEGMSVDVWPEQNLTDVISTKAFGSSYEGRRIFQLQASABRIIOA 283

Qy 242 FQKFTIPGYFFPTRNGLSKRKDKERKSLKLTERRRONAIDEGECCKEPAKDLIG 301

Db 284 FRITFIPGYWFLPKNNRRLREIEREVSKLRLGIIKGERALIKNG-----ETSGNDLLG 337

Qy 302 LMIQAK-----NVTVDIYVECKSFFPAGKQTSNLLTWITLILSMHEPQAKAR 351

Db 338 LTVSNMRBESNGKAEIGMTDEIIECKLFYFAGMETTSVLLTWITLIVSMHEPQERAR 397

Qy 352 DEVLAVCGSRDVTQKHVVKLTLSMLNESLRLPYPIATIRRAKSDVKGKYPGCT 411

Db 398 EEVLVHNG-RTTPDYSLRLKIVTWILVEVRLPVPVFLTRRYKEMELGIIKYPAEV 456

Qy 412 ELIPIAVHNDQIWMNDVNEFPARFADGVPRAKHVPVGFIPGLGYRTGIGONLAIL 471

Db 457 TLMPLFIHHDPIWKGDAEFGNPGFADGISNATKQTSFFPFMGPRICIGONFALL 516

Qy 472 QAKLTLLAVMTQRTFTHLAPTYQAHAPTYMLLYPOHGAFTTFRRL 515

Db 517 EAKKAICTIIQRFSEFELSPSYIHAPFTVITLHPHOGAQTILKXI 560

RESULT 5

US-10-425-114-59350

Sequence 59350, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jindong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(5313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 59350

LENGTH: 562

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: UC-OSFLCYP143F07_FLI.pep

US-10-425-114-59350

Query Match 42.7%; Score 1165.5; DB 12; Length 562;

Best Local Similarity 44.1%; Pred. No. 3.7e-108;

Matches 231; Conservative 97; Mismatches 175; Indels 21; Gaps 7;

Qy 5 SSSMFIPKVLVSLVSL-VIVKMSLMMRPRIEHSKQIRGPPYHFIQNKELY 63

Db 47 ASQMTLAAAAAVALMLAVSTLEWMTPRLEBALRAGQIRNRLTGDVENV 106

Qy 64 GMLKASHPMPSS-HNLPRLVSYHMRKIYGAFTLVFGPTFRITVADPLIREIFS 122

Db 107 RLNRKARKPLPGCHDIIIRVLPMFSKAVEHGKPSFTWPGTPRVMISDESIREVWS 166

Qy 123 -KSEFYKRNHAPLVKQLEBDGSLSKGKMAHHRKIISPTHEMLKLLVAVLSVTD 181

Db 167 NKFGHYGKPKPTRLGKLL-ASGVASYEGKMAHHRILNPAFHHEKIKMLVPSNCTE 225

Qy 182 MYDKSDKLSNGEVEVDYVEMFQILTEDVISRTAGSSYEDGRAVFLQAOQMLCAEA 241

Db 226 MYRWENMSIEGMSVDVWPEQNLTDVISTKAFGSSYEGRRIFQLQASABRIIOA 285

Qy 242 FQKFTIPGYFFPTRNGLSKRKDKERKSLKLTERRRONAIDEGECCKEPAKDLIG 301

Db 286 FRITFIPGYWFLPKNNRRLREIEREVSKLRLGIIKGERALIKNG-----ETSGNDLLG 339

Qy 302 LMIQAK-----NVTVDIYVECKSFFPAGKQTSNLLTWITLILSMHEPQAKAR 351

Db 340 LTVSNMRBESNGKAEIGMTDEIIECKLFYFAGMETTSVLLTWITLIVSMHEPQERAR 399

Qy 352 DEVLAVCGSRDVTQKHVVKLTLSMLNESLRLPYPIATIRRAKSDVKGKYPGCT 411

Db 400 EEVLVHNG-RTTPDYSLRLKIVTWILVEVRLPVPVFLTRRYKEMELGIIKYPAEV 458

Qy 412 ELIPIAVHNDQIWMNDVNEFPARFADGVPRAKHVPVGFIPGLGYRTGIGONLAIL 471

Db 459 TLMPLFIHHDPIWKGDAEFGNPGFADGISNATKQTSFFPFMGPRICIGONFALL 518

Qy 472 QAKLTLLAVMTQRTFTHLAPTYQAHAPTYMLLYPOHGAFTTFRRL 515

Db 519 EAKKAICTIIQRFSEFELSPSYIHAPFTVITLHPHOGAQTILKXI 562

RESULT 6

US-10-425-114-65616

Sequence 65616, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 65616
LENGTH: 547
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700100628_F11.pep
US-10-425-114-65616

Query Match 42.3%; Score 115.5; DB 12; Length 547;
Best Local Similarity 42.3%; Pred. No. 3.6e-107;
Matches 223; Conservative 104; Mismatches 181; Indels 19; Gaps 5;

QY 1 MEESSWIFPRVLTSLVTVKNSLIMPRPKIEHFSKQIGRPYHFFIGNVGMML-
DB MLREASPMALAGA-VASVSLMLVMTLEMAMTMRDLRALRAQIGKTRRLFTGDLR 86
QY 61 ELVGMMKASSHPMPS-HNLLPRLSFYHMRKYGATFLVWGPFRLTVADPDLRE 119
DB 87 ETARVNRERKRPPLGCHDITPRVQPMHSTIKYGLSTWFGPTPRVMIIDPELVK 146
QY 120 IFSKSEFYKNAHPLVQLEGDLGLSKGEKMAHHRKISPTFHEMLKLVPVYLKSV 179
DB 147 VLSNFGHFGKRSRRIRLLANGLVNHDEKMAHRRILNPAFHHRKIGMPVFSTCC 206
QY 180 TMYVDMKSDKLSENEVDVYEWFOILTEDVISRTAFSGSYEDGRAVFLQOQMLCA 239
DB 207 IEMIRDMNSMSESSSEIDVPEFONLTGVISRTAFSGSYGRIIFELQGLAERLI 266
QY 240 EAFQKVFIPGYRFFPRGNLKRKDKERKSLKLTERRRQNAIDEGECEKPAKDL 299
DB 267 QSVQITFTPGWFLPTKNNRBRADVAIRKILREIIIGKREKOD-----KNEITNDDL 320
QY 300 LGLMTQAK-----NTVQDIVECKSPFFAGKQTTNLLTWTTLISNHPQAK 349
DB 321 LGLLESNTROGNASLGLTTEVDVIEBCKLFYFAGKETTIVLLTWLLIVSNHPQER 380
QY 350 ARDEVLRVCGSRDVPKDHVVLKTLTSLMILNESLALYPIVATIRAKSDVKGKIPC 409
DB 381 ARDEVLSHFG-RTTDPYDLSRLKTVMTLHVLFLYPLATFLIRRYKMELOGIKYPA 439
QY 410 GTELLPIIAVHDAIWNADVNEFNAPADGVPRAAKAVGVIIPGLGVRTIGQNTA 469
DB 440 GVELLPIVFIHDDPIWGDASEFNPERFANGISSARHQAAPFPFGGPRICIGSFA 499
QY 470 ILQAKTLAVMIOFTFHLAPTYQHAPTYMLLYPOHGAFTFRRLT 516
DB 500 LLEAKMTLCTTIORFSELSPSYTHAPTYVITLHQRQAKIRKKLS 546

RESULT 7
US-10-424-599-183328
Sequence 183328, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 183328
LENGTH: 508
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_136559C.1.pep
US-10-424-599-183328

Query Match 41.7%; Score 1138.5; DB 12; Length 508;
Best Local Similarity 44.6%; Pred. No. 1.7e-105;
Matches 231; Conservative 95; Mismatches 163; Indels 29; Gaps 7;

QY 14 LVLSVILSVI--VKGMSLIMPRPKIEHFSKQIGRPYHFFIGNVGMML----- 67
DB 3 LLLAAVTLVFAKPFYSNFWVWTERHFKQIGGQYPIFGNTSEIRRLVAKS 62
QY 68 KASHPMPFSNHLPRVLSFYHMRKYGATFLVWGPFRLTVADPDLREIFSKSE-F 126
DB 63 EASAPPPFHHIDMGRVAPFYDMSRAYKGTFLVWGSTPRLAIVTEPDMIKELANTROE 122
QY 127 YEKNAHPLVQLEGDLGLSKGEKMAHHRKISPTFHEMLKLVPVYLKSVTDMVDRW 186
DB 123 YKVPFPQSKLFGQGLVGLBEDQNALHRRILNLAFLNELVKGWVPDVAIVTKGLSEW 182
QY 187 SDKLSENEVDVYEWFOILTEDVISRTAFSGSYEDGRAVFLQOQMLCAEAFQKVF 246
DB 183 EDQSGRGEFEIDVLRHLDSLADVISRTAFSGSYEGKILFNLQOQMLFESQAVRSY 242
QY 247 IPRGFPRTRNLSKRLDKERKSLKLTERRRQNAIDEGECEKPAKDLGLMTQA 306
DB 243 IPRGFPRTRNLSKRLDKERKSLKLTERRRQNAIDEGECEKPAKDLGLMTQA 293
QY 307 KN-----VTVQDIVECKSPFFAGKQTTNLLTWTTLISNHPQAKDEVLRYVG 359
DB 324 KNDAGGEKLGVEEIIIDCKTIYFAGKETANLLTALLLHAGQWOSVAREVHLVIG 353
QY 360 SRDVPKDHVVLKTLTSLMILNESLALYPIVATIRAKSDVKGKIPCSTELLPIIA 419
DB 364 RNRLLPAONLNDLKITMTLNETRLTPVAMLRQSKDVMGINSINPKYDPLALTA 413
QY 420 VHHQDAIWNADVNEFNAPADGV-PPAAHHPVGFIFPGGVRTIGQNTAILQAKTLA 478
DB 414 VHHREIMGEDVHNFMNRPERSPRVPR-----VRFPGGPRICIGQNTALVAKIATA 468
QY 479 VMLQFTFHLAPTYQHAPTYMLLYPOHGAFTFRRLT 516
DB 469 LIOSYSFVLSPNVMAPIFLVTLQPOYGAIIFRKIS 506

RESULT 8
US-10-424-599-199559
Sequence 199559, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 199559
LENGTH: 524
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_22226C.1.pep
US-10-424-599-199559

Qy	169	ILUSNBEVNDVYEMFOILIEDVLSRPAFGSSYDDGAVRLOQOQMLCAEAFOKFIP	248
Db	204	SLGDGSCELDVBDELQTLTGEDVLSRPAFGSSYLEGKKIFQLQAEQERMSIIDKRAVP	263
Qy	249	GYPFEPFRGNISKRLDKEIRKSLILKLIERRRONALIGEGECKEPPAKOLLGMLQIA--	306
Db	264	GYSMLPTKNNRRMQIKSEIDISIRGILGKKMQMKQGBSDK-----DDLGLLESNA	317
Qy	307	-----KNTVODIVECKSPFPAKQKOTISNLTWTWLTILSNHPEMOKARDEVL	355
Db	318	RETDGQSGPQGLTMBEWMEECKLFPYAGNETTSVLLTWTWLTILSNHPEMOKAREVL	377
Qy	356	RVCGRSDVPTQHYVVKLTLSLNLINESLRILPPIVATIRPAKSDVYKGGYKPGTEBLI	415
Db	378	GLFGRKQOPGYSRLKRTVTMTILYEVLRLPPIARSERKYYKMNVDGYTIRPGVLTLEL	437
Qy	416	PIIAVHHDQAIWGDVNEFNPARPADGVPRAKHPVGIFPGLGVPTICIGNTAILLOAKL	475
Db	438	PVLFIHHDPDIWGSAPHAFFRPERFAEGAVARASKRLAFAFPFGWGPRIICIGNFALLERAK	497
Qy	476	TLAVVIORETHLAPTQHAFTVMTLVLPOGAPITPERRL	515
Db	498	ALSMTLQRFQELAPTYTHVPRRYMTLMPMGAIQIKLRAI	537

```

RESULT 11
US-10-424-599-217648
: Sequence 217648, Application US/10424599
: Publication No. US20040031072A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kovalic David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53223)B
: CURRENT APPLICATION NUMBER: US/10/424.599
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 285684
: SEQ ID NO 217648
: LENGTH: 524
: TYPE: prt
: ORGANISM: glycine max
: FEATURE:
: OTHER INFORMATION: clone ID: PAT_MRT3847_38563C.1.pep
US-10-424-599-217648

```

Query Match	40.4%;	Score 1102;	DB 12;	Length 524;
Best Local Similarity	41.5%;	Pred. No. 8,3e-102;		
Matches 216;	Conservative 112;	Mismatches 165;	Indels 28;	Gaps 7
QY	14	LVLSEVILSYIVGMSLIMWRPKIEEHSKQIGIRPYPHFIGNVKEIVGMKLKASSHP	73	
Db	13	LIIPLVYLWCMLIKKMWLRPKLELRARAGQIGNYSLLIDTKEMTYVLMQARAQ	72	
QY	74	MPFS-----HNILPRVLSFYHWRKIVGATLVIFGCTPRVLVADDDLRELFSSSEY	127	
Db	73	QSTSSFLSKDQADAPHTTTHHHIVKFKSNKSFPEGTRKPVAVITDEQIKVEFNKIQDF	132	
QY	128	EKNELAPLVYKQBSDDGLSLKGEKNAHNRKIIISPEFHENIKLTVPVVLKSYDVMDPKMS	187	
Db	123	EKPRIISPVLKL-GSGGANLGEKMKTHKKIINPFLHEKLIKVMPLRPLFCDDMDSKHE	191	
QY	188	DKLSNGEVEVDVYEMFQILTEBDVISRFAFGSSYZDGRAVPRLOAQOQMLCAEAFQKVI	247	
Db	152	RLSSNDKSELDVWPFIQNTCDIISRTAFGSSYZDGKRIPELLKEQGLMMK-LQNAVYI	250	
QY	248	PGYRFPTFRGNLKSRKLDKIRKSLIKLTERRPQNALDGBGSECKEPPAANDLIGMTQAK	307	
Db	251	PGWMLLPFTTKMKKKIDTEIR-ALLKGVANKENAM-----KAGEVLNDDLGLMLESV	304	
QY	308	NVTYVD-----IVBECKSFFPGKQTTSNLLTWTTLLSHWPEMQAKARDEV	354	

[illegible]

```

RESULT 12
US-10-424-599-228819
; Sequence 228819, Application US/10424599
; Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yinhua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 228819
; LENGTH: 520
; TYPE: prt
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48550C.1.pep
US-10-424-599-228819

```

```

Query Match 37.5%; Score 1024; DB 12 Length 520;
Beet Local Similarity 39.4%; Pred. No. 6.1e-94;
Matches 207; Conservative 109; Mismatches 165; Indels 24; Gaps 6

QY 5 SSSNFIPTKLVLSVLSILVIVKQMSLWMPRPKILBEHFSKQIGRGPPYHFFIGNVKELG 64
DB 6 SSAALICVYVLAVALPIWALMLNLSLMLRPFRERELLRAQGFHGDYLSHHNSQNTLL 65
QY 65 MMKASHPPEFSHNILPRVLSFYHMKWIKGATFLLVWFGTEFLVADPDILREIFSKS 124
DB 66 QQHOLMSQPPPLSDVAPRVSSLHHTLIDKQKGSFFMEGRTPVILITDENQKEVFNNI 125
QY 125 EPEYKNEAHPVYKQLEGGDGLSLKGBKAAHRRKILSPFHENKLLVPEVLYKSDVMD 164
DB 126 HDFOKPFSDNVKFLFA-GILNYEEDPKAKARRLIMPAFSEKIKNNLPPFSQCHDMS 164
QY 185 KMDSKLEMBGEVADVYEWFOILTEDVYSRTAFGSGSYDGRFAVR-LQAQMLLCAEA?Q 243
DB 185 MMKGMSLSDCKRIDIMPELQNLTDVLSQNFAGSVABEKFERNRMQGYLLMAGRYK 244
QY 244 KVFIPGRFPPTGNTKSRKLDKETRFKSLKILIKRRRNALDGECECKCPAKADLLGLM 303
DB 245 N-IPILRHRTTYYKMEAIRERIDSEGLIKREKAMENG-----ETSNEDLISIL 296
QY 304 IO-----AKNVVDIYBECKSPFFAKQTTSMLLTWITLLSMHPWQAKA 350
DB 297 LESNHKEIGCHGNSRAVGMTKGVIECKFLYLAQETTSLLVWTVMLLARYBEWQARA 356
QY 351 RDEVLRVCGSRDVPKDHVYKLTLSMLNBSLELYPIVATIRASDVYLGQYKIPCG 410
DB 357 RDEYVQVFGNQN-FNIDGSKLTIWMLYEVLRKIYPTTFPSAPQCKDYKGLNLSIPAG 415
QY 411 TELIPILAVHDDALWGNVDNEFARFADGVPAACHPVGFIPLFGIQTCTGGYLA 470

```

Db 416 IRTMPLFTHHDGDIWGDADKEFKERSEGIATKQISFYFPGMGRICIGONRAL 475

Qy 471 LQAKTLAVMIQRFTHLAPYOHAPVTMLLYPOHADITERRL 515

476 MEAKTVLSILLQHFSELSFVYEHADPTVLSIQPKGAHIVAKL 520

RESULT 13

US-10-425-114-50377

Sequence 50377, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaska, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(5313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 50377

LENGTH: 458

TYPE: PR

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: LIB3060-044-A6_FLI.dep

US-10-425-114-50377

Query Match 36.2%; Score 989; DB 12; Length 458;

Best Local Similarity 41.2%; Pred No. 1.7e-90;

Matches 191; Conservative 92; Mismatches 159; Indels 22; Gaps 6;

Qy 67 LKASHMP-----FSHNLPRVLSFYHMRKIYGATFLVWFQPTFLTVADPDLR 118

Db 2 LLAHLHRRPPCGPRRLASTAPRVTFVHGRVLVBHGRTSMGSPSKITIVPDLAK 61

Qy 119 EIRF-KSEFYKNAHFLVQLBEDGLSLKGEKMAHRIISPTFMENLKLVPVYLK 177

Db 62 DVLNKGHFPEKLVPLSKVL-GSGVASHGEKVMYKRLIHAHFLSKLKMPLAST 120

Qy 178 SVTMVDMK-SDKLSENGEVEVDYEWFOILTEDVISRTAFSGSYEDGRAVRLQAQML 236

Db 121 CCEELVSRMAASIGSGSCHELDVWPELQNLTDIISRIVFSSYSEGRIRIQLOVEQAS 180

Qy 237 LCAAFQKVFIPGRFPPTSGNLKSKLDKEIRKSLIKLIERROKALIDGEGEKEPAA 296

Db 181 LMTNIRKIMIPGYMLPTANRKMNNREVSILREIIQKRIQAMEQGEGRK----- 234

Qy 297 KDILGIMTQAK-----NVTVDIYECKSPFPAGKOTTSNLTWTITLISMHEPQAKAR 351

Db 235 DDLVGLDSTNRDMDMGTIEDVIECKVFYVGMETTLVLMVWVLSMHEPQDRAR 294

Qy 352 DEVLAVCGSRDVPFKDHYVVKLTLSMLNESLRYPPIVATIRPAKSDVVLGGYKIPCGT 411

Db 295 ECVTLFGRDCKPEYDGLSKLVTTMVLVEVLRILYPRATSVROQYKEMEVGVTVBAGV 354

Qy 412 ELIPIIVHDDQAIWGDVNEFNPAFADGVPRAAHVGFIIPFGIGVTCIGONLALIQ 471

Db 355 ILLEPVLIIHDDPDIWGDADKEFKERSEGIATKQISFYFPGMGRICIGONRAL 414

Qy 472 QAKTLAVMIQRFTHLAPYOHAPVTMLLYPOHAPITERRL 515

Db 415 EAKVALLCMILQRFEGGLAPSYAHAPHTMITLHPMGAQLKRLAI 458

RESULT 14

US-10-425-114-58587

Sequence 58587, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaska, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(5313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 58587

LENGTH: 532

TYPE: PR

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: 700076168_FLI.dep

US-10-425-114-58587

Query Match 35.7%; Score 973.5; DB 12; Length 532;

Best Local Similarity 39.0%; Pred No. 7.8e-89;

Matches 204; Conservative 107; Mismatches 181; Indels 31; Gaps 9;

Qy 15 VLSVILSLVIVK-GMSIL---WVRPKIEHFSKQIGRPYHPFITGNVKELVGMKLKAS 70

Db 19 LLYVLSGLVLRAMQVLEWGLSPRLGRALRAELRGTAVFPAGAKREERLLAAR 78

Qy 71 SHMP-FSHNLPRVLSFYHMRKIYGATFLVWFQPTFLTVADPDLRIRFSK-----E 125

Db 79 AKPMSLSHAISARVPLVNAHIEHGKISVMIGTEKVLISDKLVREVTSKSKGCD 138

Qy 126 FYEKNEHAPLVKQLEDGLSLKGEKMAHRIISPTFMENLKLVPVYLKSVTMDK 185

Db 139 FQPKPLSHRIK-LIQGLTVHEGECWALHRIKISAFVYKLEMLPAFTACTGELMGR 197

Qy 186 WSKLSENGEVEVDYEWFOILTEDVISTAGSSYEDGRAVRLQAQMLCAAFQKV 245

Db 198 WESMGCDGAREIDVWPEQDITGDAISPAAGSSLSIBERRILFQAQIKI-ANEMTVL 256

Qy 246 FPGYRFPPTGRNLKSKLDKEIRKSLIKLIERROKALIDGEGEKEPAAIDLGLMIQ 305

Db 257 YIEGYTLTPKPRRKIKARAEVALLRGIITKRETMANGVAD-----SDLIGKLMQ 310

Qy 306 AK-----NVTVDIYECKSPFPAGKOTTSNLTWTITLISMHEPQAKAR 352

Db 311 SNVEAHGGGSSKPMWMDIIGELKLFYFAGMETTSALMTITLISMHEPQDRAR 370

Qy 353 EVLAVCGSRDVPFKDHYVVKLTLSMLNESLRYPPIVATIRPAKSDVVLGGYKIPCGT 412

Db 371 EVLEVFQKNE-PVINGNSLKVTTMVLVEVLRILYPRATSVROQYKEMEVGVTVBAGV 429

Qy 413 LLIPIIVHDDQAIWGDVNEFNPAFADGVPRAAHVGFIIPFGIGVTCIGONLALIQ 472

Db 430 ILIPIIIVHDDPDIWGDADKEFKERSEGIATKQISFYFPGMGRICIGONRAL 489

Qy 473 AKTLAVMIQRFTHLAPYOHAPVTMLLYPOHAPITERRL 515

Db 490 AKVALAMLQRFSGLSPTTHAPFSVSTVQPBHQAQIVYKCI 532

RESULT 15

US-10-425-114-61220

Sequence 61220, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaska, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5313) B
 CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 61220
 LENGTH: 540
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: LIB143-011-G1_F11.pcp
 US-10-425-114-61220

Query Match 35.4%; Score 967; DB 12; Length 540;
 Best Local Similarity 37.7%; Pred. No. 3.6e-88;
 Matches 205; Conservative 106; Mismatches 181; Indels 52; Gaps 9;

QY 7 SWPIPKVLVSLVSLVYKMSLM-----WRPKIEHFSKQIGRPYHPTGN 58
 DB 13 SMAFPAYLLAALLVATR---LWDYFVRLVWRPVAITKGFEGHIGSPYRFFKGC 68
 QY 59 VKELVGMMLKASHPMF-SHNLIPVLSFYHHMKIYGATLVNFGPTFRLTVADPDI 117
 DB 69 NEIRSMKERTDGLVDVGDHKLPRAPHYLEMRQYGPFLYWGQARICIDYELA 128
 QY 118 REIF-SKSEFEYKKEAHPVYKOLEGDLSTLGEKVAHHRKIISPTFHENIKLVPVL 176
 DB 129 KQISSKSGHFRVKNDAFTLLALVGKLGFMESDWRHRVRVNPFTIDKL---IVT 184
 QY 177 KSVTDMVDKSDKL-----SENGEVVDVYEMFQITLTVISRTAFSGSYEDGRAVRL 230
 DB 185 ETMLDFADSMAGELEASQENGETQVDYKHSPLVDNLAVALFGSYTLGKQVEA 244
 QY 231 QAQMMLCAEAFQVFIPIGYRFPFTGNLKRKLDEIRKSLKLIERRRONAIDGEBE 290
 DB 245 QTELGITMAFLDVPPIFGSKYLPQTANRRKMLLETKSLTRIIQPRLASGEGH--- 300
 QY 291 CKEPANDLGLMTQA-----KNTVQDIVECKSPFPAKQITSNLL 333
 DB 301 -----NDLGVMLDSCLETQGGKQOQVDRRLSLSMEEIIECKLFFAGHENTALL 354
 QY 334 TWTITLSMHPWQAKARDEVLRVCGSRDVPYKDVVYKLTLSMILNESLRLYPIVAT 393
 DB 355 TWSVYLSTIPEWQERLKEVLRFG-RDAPNPAALNRKEMTVLFETRLYSPALFMQ 413
 QY 394 RRAKSDVLAGYKIPCGTELLPIIAVHDQAIWGNVNEFNPAFADGVPPAAKPVGF 453
 DB 414 RKTADMTVGPIKLPKGTALVPIPIHNRDQAWGDADDEFSPRFANGITGAAPVHGL 473
 QY 454 IPIGIVATCTGGQALAIQALTLAVMIQRTFHLAFTYQHAPTVLMLYPOHAPITFR 513
 DB 474 LAFSGPSCIGQNLMSLEAKSTLALMKRKFSPALSPDIYAHAPVDLFTLAKKGLPVILR 533
 QY 514 RLTN 517
 DB 534 PLDN 537

Search completed: May 19, 2004, 10:33:19
 Job time : 50 secs